

00 113

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

#5-  
J1046 U.S. PRO  
09/838994  
04/20/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application:

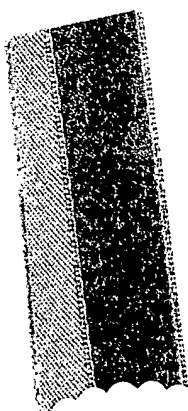
2000年 4月21日

出願番号  
Application Number:

特願2000-121853

出願人  
Applicant(s):

インターナショナル・ビジネス・マシーンズ・コーポレーション

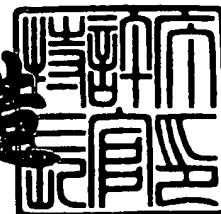


CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年12月 8日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2000-3102336

【書類名】 特許願

【整理番号】 JP9000113

【提出日】 平成12年 4月21日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

【住所又は居所】 神奈川県藤沢市桐原町1番地 日本アイ・ビー・エム株式会社 藤沢事業所内

【氏名】 山田 秀二

【特許出願人】

【識別番号】 390009531

【氏名又は名称】 インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

【識別番号】 100086243

【弁理士】

【氏名又は名称】 坂口 博

【復代理人】

【識別番号】 100112520

【弁理士】

【氏名又は名称】 林 茂則

【電話番号】 046-277-0540

【選任した代理人】

【識別番号】 100091568

【弁理士】

【氏名又は名称】 市位 嘉宏

【選任した代理人】

【識別番号】 100106699

【弁理士】

【氏名又は名称】 渡部 弘道

【選任した復代理人】

【識別番号】 100110607

【弁理士】

【氏名又は名称】 間山 進也

【選任した復代理人】

【識別番号】 100098121

【弁理士】

【氏名又は名称】 間山 世津子

【手数料の表示】

【予納台帳番号】 091156

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9706050

【包括委任状番号】 9704733

【包括委任状番号】 0004480

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データの書き込み方法およびディスクドライブ装置

【特許請求の範囲】

【請求項 1】 外部装置から書き込み要求がなされた書き込みデータを記録媒体に書き込むデータの書き込み方法であって、

書き込みコマンドを受け取るステップと、

前記書き込みコマンドに先行する先行書き込みコマンドに対応する書き込みデータに加えて、前記書き込みコマンドに対応する書き込みデータを一時記憶領域に記録するステップと、

前記一時記憶領域のフリースペースが、前記書き込みデータの一単位に相当する記憶スペースを越える所定の基準値より大きい第 1 条件を判断するステップと、を含み、

前記第 1 条件を満足する場合には、前記書き込みコマンドに対するコマンドコンプリート信号を前記外部装置に返信するデータの書き込み方法。

【請求項 2】 前記一時記憶領域に記録されている書き込みデータを前記記録媒体に書き込む指示を発するデータ書き込み指示ステップをさらに含み、

前記データ書き込み指示ステップの後に前記第 1 条件を判断して、前記第 1 条件を満足しない場合には、前記データ書き込み指示ステップに戻る請求項 1 記載のデータの書き込み方法。

【請求項 3】 前記先行書き込みコマンドと前記書き込みコマンドとの各論理的ブロックアドレスを比較して、前記書き込みコマンドの論理的ブロックアドレスの位置が前記先行書き込みコマンドの論理的ブロックアドレスの位置に対して重複または所定の離間距離以下である第 2 条件を判断するステップをさらに含み、

前記第 1 条件または前記第 2 条件の何れかの条件を満足する場合には、前記書き込みコマンドに対するコマンドコンプリート信号を前記外部装置に返信する請求項 1 記載のデータの書き込み方法。

【請求項 4】 前記一時記憶領域に記録されている書き込みデータを前記記録媒体に書き込む指示を発するデータ書き込み指示ステップをさらに含み、

前記データ書き込み指示ステップの後に前記第 1 条件を判断して、前記第 1 条件および前記第 2 条件の何れの条件をも満足しない場合には、前記データ書き込み指示ステップに戻る請求項 3 記載のデータの書き込み方法。

【請求項 5】 前記第 2 条件を満足する場合には、前記書き込みデータを前記記録媒体に書き込むステップにおいて、前記先行書き込みコマンドに対応する書き込みデータと前記書き込みコマンドに対応する書き込みデータとを連続して前記記録媒体に書き込む請求項 4 記載のデータの書き込み方法。

【請求項 6】 ランダムアクセス可能なディスク状記録媒体と、  
外部装置から送信された書き込みコマンドを実行するまでこれを保持するコマンド保持手段と、  
前記書き込みコマンドに対応する書き込みデータを一時記憶する記憶手段と、  
前記書き込みデータを前記ディスク状記録媒体に書き込むよう指示する書き込み指示手段と、

前記記憶手段のフリースペースが、前記書き込みデータの一単位に相当する記憶スペースを越える所定の基準値より大きい第 1 条件を判断する論理を備えた第 1 判断手段と、を有し、

前記第 1 判断手段により前記第 1 条件を満足すると判断された場合には、前記新規書き込みコマンドに対するコマンドコンプリート信号を前記外部装置に返信するディスクドライブ装置。

【請求項 7】 前記第 1 判断手段により前記第 1 条件を満足しないと判断された場合には、前記第 1 条件が満足されるようになるまで前記コマンドコンプリート信号の送信を待機する請求項 6 記載のディスクドライブ装置。

【請求項 8】 前記コマンド保持手段に保持された既保持書き込みコマンドと前記外部装置から新たに送信された新規書き込みコマンドとの間で論理的ブロックアドレスの重複部分がある条件、または、前記既保持書き込みコマンドと前記新規書き込みコマンドとが論理的ブロックアドレス上で規定の離間距離以下である条件、の何れかの条件を満足する第 2 条件を判断する論理を備えた第 2 判断手段をさらに有し、

前記第 1 判断手段により前記第 1 条件を満足すると判断された場合、または、

前記第 2 判断手段により前記第 2 条件を満足すると判断された場合には、前記新規書き込みコマンドに対するコマンドコンプリート信号を前記外部装置に返信する請求項 6 記載のディスクドライブ装置。

【請求項 9】 前記第 1 判断手段により前記第 1 条件を満足しないと判断され、かつ、前記第 2 判断手段により前記第 2 条件を満足しないと判断された場合には、前記第 1 条件が満足されるようになるまで前記コマンドコンプリート信号の送信を待機する請求項 8 記載のディスクドライブ装置。

【請求項 10】 前記第 2 判断手段により前記第 2 条件を満足すると判断された場合には、前記書き込み指示手段において、前記既保持書き込みコマンドと前記新規書き込みコマンドとの 2 つのコマンドに対応する書き込みデータを連続して前記ディスク状記録媒体に書き込むよう指示する請求項 9 記載のディスクドライブ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ディスクドライブ装置におけるデータの書き込みに関し、特にデータの書き込み待ち時間の低減に適用して有効な技術に関する。

【0002】

【従来の技術】

コンピュータの外部記憶装置としてのハードディスクドライブ装置に対してホストコンピュータから複数の書き込み要求がなされた場合に、実際にハードディスクドライブ装置にデータの書き込みが実行され終了するまでには相当の時間が必要である。したがって、ハードディスクドライブ装置は、書き込み要求に関するコマンドを受け取ったことをホストコンピュータに通知する一方、実行待ちの書き込みコマンドを保持することになる。

【0003】

書き込み対象のデータは、たとえばその先頭アドレスと後尾アドレスとが論理的に連続したリングキャッシュメモリに記録される。リングキャッシュメモリへのデータの記録は、ホストコンピュータからの書き込み要求の順に行われる。キ

キャッシュメモリ内の書き込みデータを実際にハードディスクドライブ装置の媒体に書き込む順番には二種類の方法がある。第1に、ホストからの書き込み要求順、つまりキャッシュに記録された順にデータを読み出し、媒体に記録する方法である。第2に、書き込みコマンドに含まれる媒体のセクタ情報を参照して書き込みの順番を再配列する方法である。

## 【0004】

前記第1の方法では、データが書き込まれる媒体のセクタが書き込み順に対してランダムであるため、ヘッドのシーク動作に時間がかかり、書き込み速度が制限される。このため、ヘッドのシーク時間が最短になるように書き込みの順番を再配列する第2の方法が多く用いられる。第2の方法の一例として、RPO (Rotational Position Optimization) を用いる方法がある。

## 【0005】

すなわち、実行待ちの書き込みコマンドが複数存在する場合には、RPOによって順次実行するコマンドを選択することが行われている。RPOは実行待ちのコマンドを実行したときに、磁気ディスク上の目標となるトラックへのシークを開始してから当該トラックに到達するまでのシーク時間と、到達後に目標のセクタにアクセスを開始するまでのディスク回転待ち時間とを予想する。そして、このシーク時間および回転待ち時間の合計が最短となる実行待ちコマンドを、次に実行するコマンドとして選択する手法である。

## 【0006】

RPOによりコマンドの実行時間を短縮することができるため、ハードディスクドライブ装置のパフォーマンスが向上する。しかし、RPOによるコマンドの実行時間をより一層短縮してハードディスクドライブ装置のパフォーマンスを向上することが望ましいことは言うまでもない。

## 【0007】

パフォーマンスを向上する一つの手法として、本願と同一の出願人にかかる先願（特願2000-019379）の技術がある。すなわち、書き込み実行待ち状態にある先行書き込みコマンドの論理的ブロックアドレスと新たに書き込み要求がなされた後続書き込みコマンドの論理的ブロックアドレスを比較して、前記

先行書き込みコマンドと前記後続書き込みコマンドとの論理的ブロックアドレス上の位置関係を判断し、この位置関係判断により前記先行書き込みコマンドと前記後続書き込みコマンドとが論理的ブロックアドレス上で離間していると判断された場合には、その離間距離が規定の離間距離以下であるか否かを判断し、前記離間距離判断により前記先行書き込みコマンドと前記後続書き込みコマンドとの論理的ブロックアドレス上の離間距離が規定の距離以下であると判断された場合には、前記先行書き込みコマンドに対応する書き込みデータと前記後続書き込みコマンドに対応する書き込みデータとを連続して前記記録媒体に書き込むデータ書き込み方法を採用できる。

## 【 0 0 0 8 】

このようなデータ書き込み方法では、実際のデータの書き込み順序は、論理的ブロックアドレス上の離間距離が短くなるように書き込まれるので、アームのシーク動作が最短距離で行われ、書き込みのための待ち時間を最短にすることができる。これにより、書き込みの効率化を図り、ハードディスクドライブ装置のパフォーマンスを向上できる。

## 【 0 0 0 9 】

## 【発明が解決しようとする課題】

しかし、前記特願 2 0 0 0 - 0 1 9 3 7 9 の方法では、以下のような問題がある。すなわち、書き込みデータは、書き込み要求の順にキャッシュメモリに記録されるため、最初に記録されたデータが書き込まれるまではキャッシュメモリが開放されない。キャッシュメモリがフルに使用された場合には書き込みコマンドを受け入れることができず、大きな待ち時間を発生させることになる。つまり、前記方法ではキャッシュに書き込まれた順序（つまり書き込み要求の順序）に関係なく、論理的ブロックアドレスが近い順にキャッシュからデータが読み出されて書き込みが行われる。このため、たとえば、媒体に実際に書き込まれたデータが記録されているキャッシュ領域（ページ）の前のページのデータが既に媒体に書き込まれているとは限らない。リングキャッシュメモリの利用を前提とする前記方法では、メモリの先頭アドレスを示すポインタは、そのポインタの位置するページが書き込まれてポインタが次ページに移動するので、ポインタが指し示す



アドレス（先頭アドレス）のデータが媒体に書き込まれるまでは、先頭アドレス以降のキャッシュメモリは開放されないこととなる。このような場合、仮に複数の書き込みコマンドのうち、先頭アドレスまたはこれに近いアドレスのデータが最後まで媒体に書き込まれないときには、最後までキャッシュメモリは開放されず、外部装置からの書き込み要求は、最後のデータが書き込まれるまで待つて受け入れることになる。つまり、書き込みの待ち時間が長くなり、パフォーマンスの向上を阻害する場合が生じる。

#### 【0010】

本発明の目的は、キャッシュメモリ（セクタバッファ）が占有される状態を回避し、書き込み時の待ち時間を短縮してハードディスクドライブ装置のパフォーマンスを向上することにある。

#### 【0011】

##### 【課題を解決するための手段】

本願発明の概要を説明すれば以下の通りである。すなわち、本願発明のデータの書き込み方法は、外部装置から書き込み要求がなされた書き込みデータを記録媒体に書き込むデータの書き込み方法であって、書き込みコマンドを受け取るステップと、前記書き込みコマンドに先行する先行書き込みコマンドに対応する書き込みデータに加えて、前記書き込みコマンドに対応する書き込みデータを一時記憶領域に記録するステップと、前記一時記憶領域のフリースペースが、前記書き込みデータの一単位に相当する記憶スペースを越える所定の基準値より大きい第1条件を判断するステップとを含み、前記第1条件を満足する場合には、前記書き込みコマンドに対するコマンドコンプリート信号を前記外部装置に返信する。

#### 【0012】

また、本願発明のディスクドライブ装置は、ランダムアクセス可能なディスク状記録媒体と、外部装置から送信された書き込みコマンドを実行するまでこれを保持するコマンド保持手段と、前記書き込みコマンドに対応する書き込みデータを一時記憶する記憶手段と、前記書き込みデータを前記ディスク状記録媒体に書き込むよう指示する書き込み指示手段と、前記記憶手段のフリースペースが、前

記書き込みデータの一単位に相当する記憶スペースを越える所定の基準値より大きい第1条件を判断する論理を備えた第1判断手段とを有し、前記第1判断手段により前記第1条件を満足すると判断された場合には、前記新規書き込みコマンドに対するコマンドコンプリート信号を前記外部装置に返信するものである。

【0013】

このようなデータ書き込み方法またはディスクドライブ装置によれば、前記第1条件を判断して、一時記憶領域（バッファメモリ）に十分な量のフリースペースが存在するかを判断できる。これにより、新規書き込みコマンドに対するコマンドコンプリートを外部装置に直ちに返送できる。逆に、バッファメモリに十分な空き領域が存在しない場合には、新規書き込みコマンドに対するコマンドコンプリートの返送を待機する。外部装置はコマンドコンプリート信号を受けない限り、次の書き込みコマンドをディスク装置に送信することがないので、ディスク装置では、実際の記録媒体への書き込みがある程度終了してバッファメモリに空き領域が開放されるまで次の書き込みコマンドを受け取る事がない。この結果、バッファメモリの空き領域が存在しないことに起因する大きな待ち時間の発生を事前に防止することができ、ディスクドライブ装置のパフォーマンスを向上できる。なお、本発明における空き領域の基準値は、従来マルチプルコマンド等で指定される最小限の空き領域を越えるバッファ空き領域を確保するためのものである。従来以上に大きなバッファ空き領域を確保することにより、書き込み順序の再配列が行われてごく初期に受け入れた書き込みコマンドの実行が遅れるような状況が発生しても、バッファメモリの占有に起因する大きな待ち時間の発生を抑制することを目的にしている。

【0014】

なお、前記データ書き込み方法またはディスクドライブ装置において、前記先行書き込みコマンドと前記書き込みコマンドとの各論理的ブロックアドレスを比較して、前記書き込みコマンドの論理的ブロックアドレスの位置が前記先行書き込みコマンドの論理的ブロックアドレスの位置に対して重複または所定の離間距離以下である第2条件を判断するステップまたは手段をさらに含み、前記第1条件または前記第2条件の何れかの条件を満足する場合には、前記書き込みコマン

ドに対するコマンドコンプリート信号を前記外部装置に返信するように構成することができる。

【0015】

このような場合、前記第2条件を判断して、新たに受け取られる書き込みコマンド（新規書き込みコマンド）が、既に一時記憶領域（バッファメモリ）に記録されている書き込みコマンドに関連するコマンドであるか否かを判断できる。これにより、前記第1条件の判断と併せて、新規書き込みコマンドが、既に保持されている書き込み待ち状態の書き込みコマンドに関連し（つまり第2条件を満足し）、あるいは、バッファメモリに十分な空き領域が存在する（つまり第1条件を満足する）場合には、新規書き込みコマンドに対するコマンドコンプリートを外部装置に直ちに返送できる。逆に、新規書き込みコマンドが書き込み待ち状態の既保持書き込みコマンドに関連せず、かつ、バッファメモリに十分な空き領域が存在しない場合には、新規書き込みコマンドに対するコマンドコンプリートの返送を待機する。コマンドコンプリート信号を受けとらない外部装置の動作は前記と同様である。この結果、バッファメモリの空き領域が存在しないことに起因する大きな待ち時間の発生を事前に防止することができ、ディスクドライブ装置のパフォーマンスを向上できる。特に、新規書き込みコマンドと既保持書き込みコマンドとの関連性がある（ヒットステータスにおいてノーヒットでない）と判断されるときには、書き込みコマンド間での書き込み順序の再配列を行うメリットが大きい一方、ヒット率が高いような場合には初期に受け入れられたコマンドの実行が後に配置される確率も高くなるので、バッファメモリに余裕を持たせる意義が大きい。

【0016】

なお、前記第2条件を満足する場合には、先行書き込みコマンド（既保持書き込みコマンド）に対応する書き込みデータと前記書き込みコマンド（新規書き込みコマンド）に対応する書き込みデータとを連続して前記記録媒体（ディスク状記録媒体）に書き込むことができる。これにより、互いに関連する書き込みコマンドで一つのコマンド塊を形成し、このコマンド塊内での書き込みシーケンスをヘッドアームのシーク時間が最短になるように最適化し、書き込み時間の短縮を

図ってディスク装置の書き込み性能を向上できる。

【0017】

【発明の実施の形態】

以下、本発明を実施の形態に基づき説明する。

【0018】

図1は、本発明を実施するコンピュータの外部記憶装置としてのHDD (Hard Disk Drive) 装置10の概略構成を示す図である。

【0019】

同図に示されるように、HDD装置10は、記憶媒体としてデータを記憶する磁気ディスク21、磁気ディスク21を回転駆動するスピンドルモータ22を備えている。磁気ヘッド23は、磁気ディスク21へのデータの記録・再生(書き込み・読み出し)を行う。ヘッドアーム24は、その先端に磁気ヘッド23を備え、磁気ディスク21の記録面の上空を移動している。また、アクチュエータ25は、ヘッドアーム24を保持すると共にヘッドアーム24を回転駆動させている。これらによって、磁気ヘッド23は、磁気ディスク21の略半径方向を移動し、磁気ディスク21の記録面に対してランダムにアクセスできるように構成されている。

【0020】

磁気ディスク21、スピンドルモータ22、磁気ヘッド23およびアクチュエータ25から構成される駆動機構は、制御回路30によって制御されている。この制御回路30は、HDC(Hard Disk Controller)31、ROM32、RAM33およびホストI/F(インターフェース)34を備えており、これらはバス35を介して相互に接続されている。

【0021】

HDC31は、HDD装置10の内部コントローラとして、ROM32に記憶された制御プログラムと制御データに従ってHDD装置10の全体を制御している。このHDC31は、サーボ制御やデータの書き込み・読み出し時の制御のための演算処理を実行している。これによって、スピンドルモータ22やアクチュエータ25を駆動し、磁気ヘッド23の記録用ヘッドおよび再生用ヘッドを用い

て書き込み・読み出しを実行している。また、HDC 3 1 は、記憶媒体としての磁気ディスク 2 1 に記録されるデータの一部を RAM 3 3 に保持し、また、磁気ディスク 2 1 に格納されているデータの一部を先読みして RAM 3 3 に保持するための制御を行っている。

## 【 0 0 2 2 】

ROM 3 2 には、HDC 3 1 により実行される制御プログラムおよびこの制御プログラムにより使用される制御データが格納されている。

## 【 0 0 2 3 】

RAM 3 3 は、磁気ディスク 2 1 に記録される書き込みデータを一時的に記憶すると共に、磁気ディスク 2 1 から読み出された読み出しデータを一時的に記憶するためのキャッシュメモリとしての機能を有する。キャッシュメモリは、物理的な開始アドレスと終端アドレスが論理的に連続して構成され、先頭アドレスがポインタにより指示されるリングキャッシュメモリが特に例示される。キャッシュメモリは、磁気ディスク 2 1 のセクタに対応するページで論理的にフォーマットされ、ページ単位でキャッシュメモリからデータが読み出されて磁気ディスク 2 1 のセクタに書き込まれ、逆に、磁気ディスク 2 1 からセクタ単位で読み出されたデータは所定のページに一時記録される。

## 【 0 0 2 4 】

また、ホスト I / F 3 4 は、ホスト 4 0 との間でデータやコマンドを送受するインターフェイス回路である。

## 【 0 0 2 5 】

図 2 は、HDC 3 1 の主要機能ブロックを示している。

## 【 0 0 2 6 】

HDC 3 1 は、MPU（マイクロ・プロセッシング・ユニット）4 1、MM（メモリ・マネージャ）4 2、RAM 4 3 および SG 4 4（セグメント・ハンドラ）を備えている。

## 【 0 0 2 7 】

RAM 4 3 には、図 1 6 に示す構成のテーブルおよび図 1 7 に示す構成のライト・セグメントがテーブル形式で格納されている。なお、図 1 6 のテーブルは、

後述する、NO OF HIT(OVERLAY HIT (オーバーレイ・ヒット) の数)、ヘッド・ベスト・スキップ・レングス (HEAD BEST SKIP LENGTH)、ヘッド・ベスト・スキップ・セグメント・ID (HEAD BEST SKIP SEGMENT ID)、テール・ベスト・スキップ・レングス (TAIL BEST SKIP LENGTH)、テール・ベスト・スキップ・セグメント・ID (TAIL BEST SKIP SEGMENT ID) を記憶する。また、図 1 7 のライト・セグメントは、4 ワードから構成されており、そのうちの 1 ワードおよび 2 ワード目を LBA が使用し、3 ワード目をブロック長 (LENGTH) が、また 4 ワード目がヒット・ステータス (HIT STATUS) が使用する構成となっている。

## 【 0 0 2 8 】

SG 4 4 は、MM 4 2 を通じて RAM 4 3 にアクセスして RAM 4 3 内のテーブルのサーチを行う。また、SG 4 4 は後述するスキップ・レングス (SKIP LENGTH) を一時的に格納するレジスタを備えている。

## 【 0 0 2 9 】

本実施の形態では、ホスト 4 0 からすでに送信された先行書き込みコマンド (既保持書き込みコマンド) と新たにホスト 4 0 から送信された後続書き込みコマンド (新規書き込みコマンド) との、論理的ブロックアドレス (以下、LBA) 上の位置関係を判断する。また、書き込みデータが一時記録されるバッファメモリ (RAM 3 3) の空き領域 (フリースペース) を判断する。これらの判断処理を制御する論理手段は、ROM 3 2 に記録されたプログラムコードにより HDC 3 1 内の MPU 4 1 で実現されてもよく、あるいは SG 4 4 または RAM 3 3 内の回路構成によりハード的に実現されても良い。なお、LBA に関する情報、つまり当該書き込みコマンドについての先端および後端の LBA およびブロック長は書き込みコマンド、書き込みデータとともにホスト 4 0 から送信される。この書き込みデータは一旦 RAM 3 3 に記憶され、また、LBA に関する情報は、図 1 7 に示したライト・セグメントに記憶される。

## 【 0 0 3 0 】

本実施の形態では、先行書き込みコマンド (既保持書き込みコマンド) のことをバッファデータと称し、後続書き込みコマンド (新規書き込みコマンド) のことを単に書き込みコマンドと称することとする。また、本実施の形態では、LB

A上の位置関係を、単に、関係、あるいは位置関係と言うことがある。

【0031】

図22は、本実施の形態のデータの書き込み方法を示したフローチャートである。外部装置40から送信された書き込みコマンドを、ホストI/F34を介して受け取る(S901)。受け取られた書き込みコマンドは、HDC31内のMPU41で解読され、LBAおよびブロック長に関する情報がRAM43に転送される。

【0032】

次に、書き込みコマンドに関連する書き込みデータを受け取る(S902)。受け取られた書き込みデータは、RAM33に記録される。なお、RAM33への記録形態は、後に詳述する。

【0033】

次に、ヒットステータス(Hit Status)がチェックされ、書き込みコマンドの実行(つまり記録媒体21への書き込み実行)の順番に関する再配列が行われる(S903、S904)。ヒットステータスは、新規に受け入れられた書き込みコマンド(すなわちS901で受け入れられた書き込みコマンド)と既にRAM43に記録されている書き込み待ち状態の書き込みコマンド(既保持書き込みコマンド)との関係を表す。ヒットステータスのチェックおよびコマンド実行順の再配列は、第1段階としてコマンドの関係を図3に示すような8つの分類の何れに該当するかをLBAに基づいて判断し、第2段階としてその関係を詳細に判断する。

【0034】

図3は、バッファデータ(既保持書き込みコマンド)と新規書き込みコマンドとの関係を示す概念図である。NO HIT-A、TAIL\_SEQ、OV\_TAIL、OV\_ALL、OV\_COVER、OV\_HEAD、HEAD\_SEQおよびNO HIT-Bの8つの分類のうちのいずれに該当するかをLBAに基づき判断する。この判断は、図4～図6に示すフローに基づき行われる。

【0035】

図3において、バッファデータの書き込み先の後端のLBAをLLBAと表記

する。また、バッファデータのブロック長をLENGTHと表記している。さらに、書き込みコマンドの書き込み先の先端のLBAをSLBAと、後端のLBAをELBAと表記する。さらにまた、書き込みコマンドのLBAのブロック長をSLENという。

#### 【0036】

図3中の#1～#8と記されているものが書き込みコマンドを示している。なお、#1～#8は書き込みコマンドを便宜上区分けするために記したものであって、ホスト40から送信された順番を示すものでないことを最初に断っておきたい。

#### 【0037】

本実施の形態においては、バッファデータに対する書き込みコマンドの位置関係を、図3に示すように8種類に分類する。各々の内容は以下の通りである。

#### <NO HIT-A>

書き込みコマンド#1は、バッファデータのLLBAより離れた位置に存在している。このようにバッファデータに対して書き込みコマンド#1が重複部分を有していない場合には、NO HIT-Aと判断する。

#### <TAIL\_SEQ>

書き込みコマンド#2のように、バッファデータの後端に書き込みコマンド#2の先端が連続的につながっている場合には、TAIL\_SEQと判断する。

#### <OV\_TAIL>

書き込みコマンド#3のように、バッファデータの後端側でバッファデータと書き込みコマンド#3が重複している場合には、OV\_TAILと判断する。

#### <OV\_ALL>

書き込みコマンド#4は、バッファデータの中に全て包含されている。このような場合にはOV\_ALLと判断する。

#### <OV\_COVER>

書き込みコマンド#5は、バッファデータを全て包含している。このような場合にはOV\_COVERと判断する。

#### <OV\_HEAD>



書き込みコマンド# 6のように、バッファデータの先端側でバッファデータと書き込みコマンド# 6が重複している場合には、OV\_HEADと判断する。

<HEAD\_SEQ>

書き込みコマンド# 7のように、バッファデータの先端に書き込みコマンド# 7の後端が連続的につながっている場合には、HEAD\_SEQと判断する。

<NO HIT-B>

書き込みコマンド# 8のように、バッファデータの先端側において書き込みコマンド# 8がバッファデータと離れている場合には、NO HIT-Bと判断する。

【0038】

以上の判断は、HDC 31のSG（セグメント・ハンドラ）44が、書き込みコマンドのLBA情報と、RAM 33内に記憶されているセグメント・テーブルのLBA情報を比較することにより実行される。具体的なフローを図4～図6に基づき説明する。

【0039】

図4～図6において、ステップS11において、 $LLBA < SLBA - 1$ であるか否か判断し、 $LLBA < SLBA - 1$ の場合にはステップS12に進み、そうでない場合にはステップS13に進む。なおステップS11における判断は、バッファデータの後端のLBAと書き込みコマンドの先端のLBAとの位置を比較したものである。 $LLBA < SLBA - 1$ であれば、ステップS12において、バッファデータと書き込みコマンドとの間にLBA上の重複がないことから、バッファデータに対する書き込みコマンドの関係、つまりヒット・ステータス（HIT STATUS）をNO HIT-Aと判断する。なお、「A」は、書き込みコマンドがバッファデータの後方に存在することを示している。

【0040】

ステップS13では $LLBA = SLBA - 1$ であるか否か、つまりバッファデータの後端のLBAに書き込みコマンドの先端のLBAが連続するものであるか否か判断する。連続していればステップS14においてヒット・ステータスをTAIL\_SEQと判断する。 $LLBA = SLBA - 1$ でなければステップS15

に進む（以下、図5）。

【0041】

ステップS15では、 $LENGTH < LLBA - SLBA + 1$ であるか否か判断する。そうであればステップS19に進み、そうでなければステップS16に進む。

【0042】

ステップS16では $LLBA < ELBA$ であるか否かを判断する。 $LLBA < ELBA$ であればステップS17においてヒット・ステータスをOV\_TAILと判断し、そうでなければステップS18においてヒット・ステータスをOV\_ALLと判断する。

【0043】

一方、ステップS19では $LLBA \leq ELBA$ であるか否か判断し、そうであればステップS20においてヒット・ステータスをOV\_COVERと判断する。そうでなければ、ステップS21に進む。

【0044】

ステップS21では、 $LENGTH < LLBA - ELBA$ であるか否か判断し、そうであればステップS22においてヒット・ステータスをNO\_HIT-Bと判断する。そうでなければ、ステップS23に進む（以下、図6）。

【0045】

ステップS23では、 $LENGTH = LLBA - ELBA$ であるか否か判断し、そうであればヒット・ステータスをHEAD\_SEQと判断する。一方、そうでなければヒット・ステータスをOV\_HEADと判断する。

【0046】

以上で第1段階の判断が終了する。

【0047】

次に、第2段階の処理を図7～図14に基づき説明する。

<NO\_HIT-A>

NO\_HIT-Aと判断された場合には、図7に示す判断がなされる。つまり、ステップS101においてスキップ・レンジ（SKIP\_LENGTH）が

スキップ・クライテリア (SKIP CRITERIA) より大きいかな否か判断する。なお、ここでスキップ・レンジとは、 $LLBA - SLBA + 1$ で定義される値であり、バッファデータの後端と書き込みコマンドの先端との距離を示している。

## 【0048】

スキップ・レンジがスキップ・クライテリアより大きい場合には、ステップ S102において、書き込みコマンドはバッファデータに対してNO HITであると判断する。つまり、書き込みコマンドの先端とバッファデータの後端との離間距離が規定の離間距離であるスキップ・クライテリアを超えて離れている場合にはNO HITと判断する。NO HITと判断されると、当該バッファデータと当該書き込みコマンドとの間に関連性がないものとみなして、データの書き込み実行時に連続して書き込む等の処置を行わない。

## 【0049】

スキップ・レンジがスキップ・クライテリア以下の場合には、書き込みコマンドは、ステップ S103において、バッファデータに対してテール・ヒット (TAIL HIT) であると判断する。つまり、当該バッファデータと当該書き込みコマンドとは離間しているものの、その距離が規定値以下であるために連続性のあるデータに関するものとみなしてデータの書き込み実行時に連続して書き込みを実行する対象とする。したがって、RPOによるパフォーマンス向上に有効である。この点は、NO HIT-Bと判断された場合についても同様である。

## 【0050】

テール・ヒットと判断した場合には、今までのテール・ベスト・スキップ・レンジ (TAIL BEST SKIP LENGTH) が0未満であるかな否かステップ S104にて判断する。テール・ベスト・スキップ・レンジが0未満の場合には処理が終了する。そうでなければステップ S105において、今回のスキップ・レンジがテール・ベスト・スキップ・レンジより小さいかな否か判断する。小さい場合には、ステップ S106において、今回のスキップ・レンジをテール・ベスト・スキップ・レンジとし、テール・ベスト・スキップ・レ

ングスおよびテール・ベスト・スキップ・セグメント・IDを更新する。そうでなければ、処理を終了する。

【0051】

ここで、テール・ベスト・スキップ・レンジスについて、図20を参照しつつ説明する。図20は、書き込みコマンドが2つのバッファデータのいずれに対してもテール・ヒットとなっているものとする。ところが、スキップ・レンジスは、セグメントIDが2のバッファデータのほうがセグメントIDが1のバッファデータデータよりも小さいことがわかる。この場合、当該書き込みコマンドのテール・ベスト・スキップ・レンジスはセグメントIDが2のバッファデータとのスキップ・レンジスとなり、テール・スキップ・セグメント・IDは2になる。そして、テール・ベスト・スキップ・レンジスおよびテール・スキップ・セグメント・IDは、図16に示すテーブルに記憶される。

<TAIL\_SEQ>

TAIL\_SEQと判断された場合には、図8に示す処理がなされる。

【0052】

TAIL\_SEQの場合には、ステップS201にて、ヒット・ステータスをテール・ヒットと判断する。次いで、今までのテール・ベスト・スキップ・レンジスが0未満であるか否か、ステップS202にて判断する。テール・ベスト・スキップ・レンジスが0未満の場合には処理を終了する。そうでなければステップS203において、スキップ・レンジスがテール・ベスト・スキップ・レンジスより小さいか否かを判断する。小さい場合には、ステップS204において、テール・ベスト・スキップ・レンジスおよびテール・ベスト・スキップ・セグメントIDを更新する。

<OV\_TAIL>

OV\_TAILと判断された場合には、図9に示す処理がなされる。

【0053】

まず、ステップS301にてヒット・ステータスをテール・ヒットと判断し、ついで、ステップS302にてスキップ・レンジスがテール・ベスト・スキップ・レンジスより小さいか否かを判断する。小さい場合には、ステップS303に

において、テール・ベスト・スキップ・レンジおよびテール・ベスト・スキップ・セグメントIDを更新する。そうでない場合には処理を終了する。

#### <OV\_TALL>

OV\_TALLと判断された場合には、図10および図11に示す処理がなされる。

##### 【0054】

まず、ステップS401においてLLBAとELBAとが一致するか否か判断する。一致する場合にはステップS402に進み、一致しない場合には図11のステップS408に進む。

##### 【0055】

ステップS402では、 $SLBA = LLBA - LENGTH + 1$ であるか否かを判断する。そうであればステップS403へ、そうでなければステップS405へ進む。

##### 【0056】

ステップS404では、ステップS402でSLBAとバッファデータの先端LBAとが一致すると判断した場合に、ヒット・ステータスをカバー・オーバーレイと判断し、次いでステップS404においてオーバーレイ・ヒット(OVERLAY HIT)のカウンタを1増やす。

##### 【0057】

ステップS405では、ヒット・ステータスをテール・ヒットと判断する。次いでステップS406にてスキップ・レンジがテール・ベスト・スキップ・レンジより小さいか否かを判断する。小さい場合には、ステップS407において、テール・ベスト・スキップ・レンジおよびテール・ベスト・スキップ・セグメントIDを更新する。そうでない場合には処理を終了する。

##### 【0058】

ステップS408では $SLBA = LLBA - LENGTH + 1$ であるか否か判断する。そうであればステップS409へ、そうでなければステップS412へ進む。

##### 【0059】

ステップS409ではヒット・ステータスをヘッド・ヒット (HEAD HIT) と判断する。続いて、ステップS410において、スキップ・レングス (ここでは、SLEN) がヘッド・ベスト・スキップ・レングス (HEAD BEST SKIP LENGTH) より小さいか否かを判断する。そうであればステップS411へ進み、また、そうでなければ処理を終了する。

## 【0060】

ステップS411ではヘッド・ベスト・スキップ・レングスおよびヘッド・ベスト・スキップ・セグメントID (HEAD BEST SKIP SEGMENT ID) を更新し、その後処理を終了する。

## 【0061】

S412ではヒット・ステータスをオール・オーバーレイ (ALL OVERLAY) と判断し、ついでステップS413にてオーバーレイ・ヒットのカウントを1増やして処理を終了する。なお、オーバーレイ・ヒットのカウントは、図16に示すテーブルのNO OF HITのセルにて増やされる。

## &lt;OV\_COVER&gt;

OV\_COVERと判断された場合には、図12に示すようにステップS501においてヒット・ステータスをオール・オーバーレイと判断し、次いでステップS502にてオーバーレイ・ヒットのカウントを1増やして処理を終了する。

## &lt;OV\_HEAD&gt;

OV\_HEADと判断された場合には、図13に示す処理がなされる。

## 【0062】

OV\_HEADの場合には、ステップS601でヒット・ステータスをヘッド・ヒットと判断する。次いで、ステップS602において、スキップ・レングス (この判断では、 $(LLBA - LENGTH + 1) - ELBA$ ) がヘッド・ベスト・スキップ・レングスより小さいか否かを判断する。小さい場合には、ステップS603において、ヘッド・ベスト・スキップ・レングスおよびヘッド・ベスト・スキップ・セグメントIDを更新する。そうでなければ処理を終了する。

## &lt;HEAD\_SEQ&gt;

HEAD\_SEQと判断された場合には、図14に示す処理がなされる。

## 【0063】

HEAD\_SEQの場合には、ステップS701においてヒット・ステータスをヘッド・ヒットと判断する。次いでステップS702において今までのベスト・ヒット・スキップ・レングスが負か否かを判断する。負の場合には処理を終了する。負でない場合には、ステップS703に進む。

## 【0064】

ステップS703では、スキップ・レングスがヘッド・ベスト・スキップ・レングスより小さいか否かを判断する。なお、このスキップ・レングスはゼロである。スキップ・レングスがヘッド・ベスト・スキップ・レングスより小さい場合には、ステップS704においてヘッド・ベスト・スキップ・レングスおよびヘッド・ベスト・スキップ・セグメント・IDを更新して処理を終了する。

## &lt;NO HIT-B&gt;

NO HIT-Bと判断された場合には、図15に示す処理がなされる。つまり、ステップS801において、 $(LLBA - LENGTH + 1) - ELBA$ がスキップ・クライテリアより大きいと判断する。

## 【0065】

スキップ・レングスがスキップ・クライテリアより大きい場合には、ステップS802において、ヒット・ステータスをノー・ヒット (NO HIT) と判断する。

## 【0066】

スキップ・レングスがスキップ・クライテリア以下の場合には、ステップS803において、ヒット・ステータスがヘッド・ヒットであると判断し、ステップS804に進む。

## 【0067】

ステップS804ではスキップ・レングスがヘッド・ベスト・スキップ・レングスより小さいか否かを判断する。なお、このスキップ・レングスは、 $(LLBA - LENGTH + 1) - ELBA$ で定義される。ヘッド・ベスト・スキップ・レングスがスキップ・レングス以下である場合には処理を終了する。スキップ・レングスがヘッド・ベスト・スキップ・レングスより小さい場合には、ステッ

ブ S 8 0 5 において、ヘッド・ベスト・スキップ・レングスおよびヘッド・ベスト・スキップ・セグメント I D を更新して処理を終了する。

## 【 0 0 6 8 】

以上説明したように、新規書き込みコマンドが受け入れられた時点でヒットステータスが検査され、ヒットステータスに応じて書き込みコマンドの再配列が行われる。なお、前記フローで N O H I T と判断された場合には、新規書き込みコマンドと既に保持されている書き込み待ちのコマンドとの間に関係がないとしてコマンドの再配列は行われず、それ以外のステータスでは新規コマンドと既存のコマンドとの間に何らかの関係があると判断されている。N O H I T - A 、 T A I L \_ S E Q 、 O V \_ T A I L 、 O V \_ A L L 、 O V \_ C O V E R 、 O V \_ H E A D 、 H E A D \_ S E Q および N O H I T - B と図 7 ～ 図 1 5 により判断された各ヒット・ステータスとの関係をバッファデータとの位置関係を含めて図 2 1 に示しておく。

## 【 0 0 6 9 】

以上の判断はホスト 4 0 から書き込み要求がなされる度にセグメント・ハンドラ ( S G ) がメモリ・マネージャ ( M M ) を通じてセグメント・テーブルが格納されている R A M 4 3 にアクセスすることにより実行される。セグメント・テーブルを構成するライト・セグメントは、図 1 6 に示すように 4 ワードから構成される。そのうちの 1 ワードおよび 2 ワード目を L B A が使用し、 3 ワード目をブロック長 ( L E N G T H ) が、また 4 ワード目がヒット・ステータスが使用する構成となっている。

## 【 0 0 7 0 】

この書き込みセグメントに格納されたヒット・ステータスを R P O ( R o t a t i o n a l P o s i t i o n O p t i m i z a t i o n ) に反映させ、実際の書き込みプロセスの実行を指示する ( S 9 0 5 ) 。実際の書き込みプロセスは S 9 0 1 ～ S 9 0 7 のフローとは別のプロセスで行う ( S 9 1 0 ) 。

## 【 0 0 7 1 】

S 9 1 0 のプロセスを以下に説明する。R P O は、まず実行待ちのコマンドを実行したときに、磁気ディスク 2 1 上の目標となるトラックへのシークを開始し



てから当該トラックに到達するまでのシーク時間と、到達後に目標のセクタにアクセスを開始するまでのディスク回転待ち時間とを予想する。そして、このシーク時間および回転待ち時間の合計が最短となる実行待ちコマンドを、次に実行するコマンドとして選択する手法である。

## 【 0 0 7 2 】

本実施の形態においては、前述のようにライト・セグメントに格納されたヒット・ステータスを参照することにより、複数の書き込みコマンドのうち、一連のデータとみなせる書き込みコマンドを関連づけることができる。つまり、相互に N O - H I T 以外の場合には、一連のコマンドの塊として連続的に書き込み実行の対象とすることができる。

## 【 0 0 7 3 】

本実施の形態では前述のように書き込みセグメントを 6 4 セット用意しているため、その中で以上のようなコマンド塊が複数存在する場合もある。このようにコマンド塊が複数存在した場合には、従来の R P O にしたがって、シーク時間および回転待ち時間の合計が最短となるコマンド塊から先に書き込みの実行を行う。もちろん、コマンド塊を構成しない書き込みコマンドが存在することもあるから、そのような書き込みコマンド同士、また、コマンド群との間では従来の R P O にしたがって、書き込みコマンドの実行順位が決定される。

## 【 0 0 7 4 】

以上の処理の概念について、図 1 8 を用いて説明する。

## 【 0 0 7 5 】

図 1 8 ( a ) は、ホスト 4 0 からの書き込み要求のあった複数のコマンド①～⑥を示している。各コマンドに付された番号①～⑥は、ホストから送信された書き込みの要求順位を示す。各コマンドに対応する書き込みデータは、書き込み要求順にキャッシュメモリ（バッファメモリ）に記録されるので、番号①～⑥は、キャッシュメモリのアドレス（ページ）にも対応する。また、横軸は L B A を示している。

## 【 0 0 7 6 】

図 1 8 ( a ) において、前述の L B A の比較に基づき、コマンド①とコマンド

③、また、コマンド②とコマンド⑤は各々は一連のデータ群を構成していると判断する。しかし、コマンド④、コマンド⑥は、他のコマンドとの関連がなく、ノー・ヒットと判断されるものである。したがって、コマンド①～⑥は、図 1 8 (b) に示すように A～D の 6 つのデータ群に置き換えることができる。そして、書き込みを行う際の順番を、ホスト 4 0 から送信された順番ではなく、A, B, C および D の順番となるように並び替える。

【0 0 7 7】

なお、オーバーレイ・ヒットとなったデータ群の書き込み実行を行う際の処理について図 1 9 に基づき説明する。

【0 0 7 8】

図 1 9 に示すように、先行する書き込み要求のコマンド（バッファデータ）の L B A が 1, 2 および 3、後続の書き込み要求のコマンド（書き込みコマンド）の L B A が 0 および 1 とする。バッファデータに対する書き込みコマンドのヒット・ステータスは、オーバーレイ・ヒットであり、L B A 1 のブロックが重複している。R P O に単純に従って書き込みコマンドの書き込みを実行した後にバッファデータの書き込みを実行したとすると、L B A 1 のブロックについては、書き込みコマンドのデータの上にバッファデータのデータが上書きされることになる。バッファデータのデータは書き込みコマンドのデータより古いから、前記上書きにより古いデータが新しいデータの上に書き込まれることになる。

【0 0 7 9】

そこで本実施の形態において、図 1 9 の右側のように、バッファデータのうち重複するブロックを消去（D I S C A R D）することができる。その後に、バッファデータの書き込み実行し、ついで書き込みコマンドの書き込みを実行することができる。そうすることにより、重複する部分についても新しいデータが書き込まれることになる。

【0 0 8 0】

以上のようにして書き込みコマンドが順次実行され、R A M 3 3 に記録された書き込みデータが記録媒体 2 1 に記録される（S 9 1 0）。従来技術では、キャッシュあるいはバッファ領域（R A M 3 3）に空き領域がある限り、新規書き込

みコマンドに対するコマンドコンプリートが書き込みデータをバッファ領域（RAM 3 3）に記録した段階で発行されるため、次の書き込みコマンドが次々に送られてくる。このようにして送付された書き込みコマンドは、実際の媒体への書き込みプロセス（S 9 1 0）と並行して前記の通り処理される。

## 【 0 0 8 1 】

一方、記録媒体 2 1 にデータが書き込まれた時点で、RAM 3 3 上の書き込みデータは不要になり開放されるべきである。しかし、RAM 3 3 にリングキャッシュを用いている場合には、バッファ領域は記録媒体 2 1 への書き込みにより順次開放されるわけではない。

## 【 0 0 8 2 】

たとえばキャッシュの開始アドレスポインタが図 1 8 (a)においてデータ①の先頭にあると仮定し、たとえば図 1 8 (b)において、データ群 A、B、C、D の順にデータが書き込まれるとする。データ群 A のデータ⑥が記録媒体 2 1 に書き込まれても、データ⑥が記録されていたキャッシュ領域は開放されず、キャッシュの開始アドレスポインタはデータ①の先頭であることに変化はない。データ群 B のデータ①が記録媒体 2 1 に書き込まれた時点でデータ①は不要になるのでキャッシュの開始アドレスポインタはデータ②の先頭に移動する。ここで初めてキャッシュ領域はデータ①に相当する分だけ開放されたことになる。さらに書き込みが進行してデータ③が書き込まれても、キャッシュ開始アドレスポインタの位置はデータ②の先頭であることに変化はない。次のデータ群 C のデータ④が書き込まれても同様である。書き込みがデータ群 D のデータ②にまで進行して初めて開始アドレスポインタはデータ⑤の先頭に移動する。これにより既に記録媒体 2 1 に書き込まれて不要になったデータ②③④に相当する領域が開放されることになる。つまり、初期に受け入れられた書き込みコマンドが P R O による再配列で最後の方に配置された場合には、この初期に受け入れられた書き込みコマンドに対応する書き込みデータが実際に記録媒体 2 1 に書き込まれない限り、キャッシュ領域が開放されない。データ群を構成するデータについても同様な状況が生じる。

## 【 0 0 8 3 】

バッファ領域に十分な余裕が常に存在する場合には特に問題を生じないが、新たな書き込みコマンドの受け入れによる書き込みデータの記録が不可能な程度にまでバッファ領域が占有されていると、データの上書き防止のため、新規コマンドを受け入れることはできない。この場合、新規コマンドの受け入れは、初期に受け入れられた書き込みコマンドに対応する書き込みデータが実際に記録媒体 2 1 に書き込まれてキャッシュ領域が開放されるまでコマンドコンプリートの発行を待たなければならない。

## 【 0 0 8 4 】

本実施の形態では、このような問題点の解決のために、セクタバッファの空き領域（フリースペース）をモニタし、所定の基準値よりも小さな空き領域になった場合には、受け入れた書き込みコマンドに対するコマンドコンプリート信号の発行を保留して記録媒体 2 1 への実際の書き込み（S 9 1 0）を優先させる。ただし、ヒットステータスが N O H I T 以外の場合（つまり受け取った書き込みコマンドが既保持のコマンドと関連があると判断した場合）には、バッファの空き領域にかかわらずコマンドコンプリートを発行して次のコマンドの受け取りを可能にする。図 2 2 のフローチャートで説明すれば以下の通りである。前記した実際の書き込みプロセスを指示（S 9 0 5）した後、ヒットステータスが N O H I T でない条件（第 2 条件）と、バッファのフリースペースが基準値を超える条件（第 1 条件）とを判断し（S 9 0 6）、第 1 条件または第 2 条件のいずれかを満足している場合には、S 9 0 1 で受け入れたコマンドに対するコマンドコンプリート信号を発行し（S 9 0 7）、それ以外の場合は S 9 0 5 に戻る。

## 【 0 0 8 5 】

このような判断ステップを含むことにより、バッファ領域に所定の空き領域が確保されるまでコマンドコンプリート信号の発行が保留されることになり、バッファ領域には常にある程度の空き領域が確保されることになり、バッファ領域（R A M 3 3）の占有によるパフォーマンスの低下を防止することができる。

## 【 0 0 8 6 】

本実施の形態では、単に基準値以上の空き領域をバッファに確保しようとするだけでなく、受け入れたコマンドが既保持のコマンドと関連する場合にはバッフ

アの空き領域が基準値以下でもコマンドコンプリートを発行する。これは、現に受け入れているコマンド（新規コマンド）が既存のコマンドに関連している場合（つまりNO HIT以外の場合）には、前記したヒットステータスごとの処理を行うことによりパフォーマンス向上のメリットが生かせる期待値が大きいことによる。一方NO HITの場合には前記処理によるパフォーマンスの向上は期待できず、逆にその後のコマンドの受け入れにより、RPOによる処理順序の再配列の結果、処理順序が後の方に回されれば、前記した問題を生じる可能性が高くなる。そこで、新規コマンドがNO HITであり、かつバッファ空き領域が基準値以下の時には、実際の書き込みプロセス（S910）を優先させて、バッファの空き領域を確保するようにしたものである。

## 【0087】

現コマンドに対するコマンドコンプリートが発行された後には、外部装置40は、次の書き込みコマンドを送信でき、上記と同様に処理されることとなる。

## 【0088】

本実施の形態によれば、上記した通り、バッファメモリの空き領域を常にある程度確保することができるため、バッファ領域の占有によるパフォーマンスの低下を防止できる。

## 【0089】

図23および図24はバッファ空き領域の基準値を変化させたときのベンチマークテストのスコアを相対的に示したグラフであり、図23はハイエンドユースを目的としたベンチマークの結果であり、図24はビジネスユースを目的としたベンチマークの結果である。ベンチマークテストには、Ziff Davis社製の「WinBench 99（商標）」内のディスクアクティビティテスト「WinMark 99（商標）」を用いている。WinMark 99におけるハイエンドスイート（High-End Suite）では、たとえばアドビ社製「Photoshop（商標）4.01」等の画像あるいは音響データ等大量のデータを扱うアプリケーションソフトを利用した時のパフォーマンスが測定でき、ビジネススイート（Business Suite）では、たとえばロータス社製「Word Pro（商標）97」等の一般的なアプリケーションソフトを利用した時のパフォーマンスが測定できる。このテストに用いたディスク装置のバッファサイズは30

0 (16進数) ページ (10進数では768ページ) であり、記録媒体サイズは2.5インチ、媒体枚数1枚、記録容量10GB、スピンドル回転数4200rpm、RAMサイズ4MBである。ディスク装置が取り付けられているシステムは一般的なノートブック型パーソナルコンピュータ (CPU: Pentium-II (商標)、クロック周波数366MHz、RAMサイズ128MB) である。

## 【0090】

図23に示すように、基準値が0ページの場合 (従来技術に相当する) を1として規格化すると、基準値が約450 (10進数) ページ (つまり450ページのフリースペースを確保する) になるまで基準値を増加していくと順次パフォーマンスが向上し、その後基準値を増加するに従いパフォーマンスが低下する傾向にある。つまり基準値が約450 (10進数) ページの点に極大値を持つ。パフォーマンスは、従来技術相当の条件 (基準値0) に比較して25%の向上が見られる。

## 【0091】

一方、図24に示すように、ビジネススイートの場合には、基準値が450までは従来技術相当の条件 (基準値0) に比較して大きな変化は見られないが、基準値450を超える頃からパフォーマンスが急激に低下する。

## 【0092】

結果として、本実施の形態を用い、その基準値を450程度に設定すれば、ハイエンドにおいて25%程度のパフォーマンスの向上が得られ、同条件でのビジネスユースでも性能の低下は見られない。すなわち、総合的なディスク装置の性能の向上が図れる。

## 【0093】

以上実施の形態に基づき本願発明を説明したが、本願発明は前記実施の形態に限定されずその要旨を変更しない範囲で種々変更可能であることは勿論である。たとえば、HDC31内のRAM43の機能をRAM33に持たせてもよく、逆にHDC31内のRAM43の容量を増加してRAM33の機能をRAM43に持たせてもよい。この場合RAM33は必要なくなる。また、ヒットステータス

ごとの処理において、各書き込みコマンド（セグメント）に書き込み順序を反映したチェーン情報を関連付け、このチェーン情報に従って記録媒体 21 への書き込み処理を行っても良い。

【0094】

【発明の効果】

以上説明のように、本発明によれば、キャッシュメモリ（セクタバッファ）が占有される状態を回避し、書き込み時の待ち時間を短縮してハードディスクドライブ装置のパフォーマンスを向上することができる。また、実行待ちの書き込みコマンドがある場合の書き込みの効率化を図り、ハードディスクドライブ装置のパフォーマンスを向上することができる。

【図面の簡単な説明】

【図1】 本実施の形態にかかるHDD（Hard Disk Drive）装置の概略構成を示す図である。

【図2】 本実施の形態にかかるHDDのHDC（Hard Disk Controller）の主要構成を示す図である。

【図3】 バッファデータと書き込みコマンドとの関係を説明するための図である。

【図4】 バッファデータと書き込みコマンドとの関係を判断するためのフローチャートある。

【図5】 バッファデータと書き込みコマンドとの関係を判断するためのフローチャートである。

【図6】 バッファデータと書き込みコマンドとの関係を判断するためのフローチャートである。

【図7】 バッファデータと書き込みコマンドとの関係がNO HIT-Aの場合のその後の処理を示すフローチャートである。

【図8】 バッファデータと書き込みコマンドとの関係がTAIL\_SEQの場合のその後の処理を示すフローチャートである。

【図9】 バッファデータと書き込みコマンドとの関係がOV\_TAILの場合のその後の処理を示すフローチャートである。

【図 1 0】 バッファデータと書き込みコマンドとの関係がOV\_ALLの場合のその後の処理を示すフローチャートである。

【図 1 1】 バッファデータと書き込みコマンドとの関係がOV\_ALLの場合のその後の処理を示すフローチャートである。

【図 1 2】 バッファデータと書き込みコマンドとの関係がOV\_COVERの場合のその後の処理を示すフローチャートである。

【図 1 3】 バッファデータと書き込みコマンドとの関係がOV\_HEADの場合のその後の処理を示すフローチャートである。

【図 1 4】 バッファデータと書き込みコマンドとの関係がHEAD\_SEQの場合のその後の処理を示すフローチャートである。

【図 1 5】 バッファデータと書き込みコマンドとの関係がNO\_HIT-Bの場合のその後の処理を示すフローチャートである。

【図 1 6】 RAM 4 3 に格納されるテーブルを示す図である。

【図 1 7】 ライト・セグメントの構成を示す図である。

【図 1 8】 本実施形態によるデータの書き込みを説明するための図である。

【図 1 9】 本実施形態によるオーバーレイ・ヒットの時のデータの書き込みを説明するための図である。

【図 2 0】 ベスト・テール・スキップ・レンジを説明するための図である。

【図 2 1】 図 7 ～ 図 1 5 により判断された各ヒット・ステータスとの関係をバッファデータとの位置関係を含めて示した図である。

【図 2 2】 本実施の形態のデータの書き込み方法を示したフローチャートである。

【図 2 3】 バッファ空き領域の基準値を変化させたときのハイエンドスイートベンチマークテストのスコアを相対的に示したグラフである。

【図 2 4】 バッファ空き領域の基準値を変化させたときのビジネススイートベンチマークテストのスコアを相対的に示したグラフである。

【符号の説明】

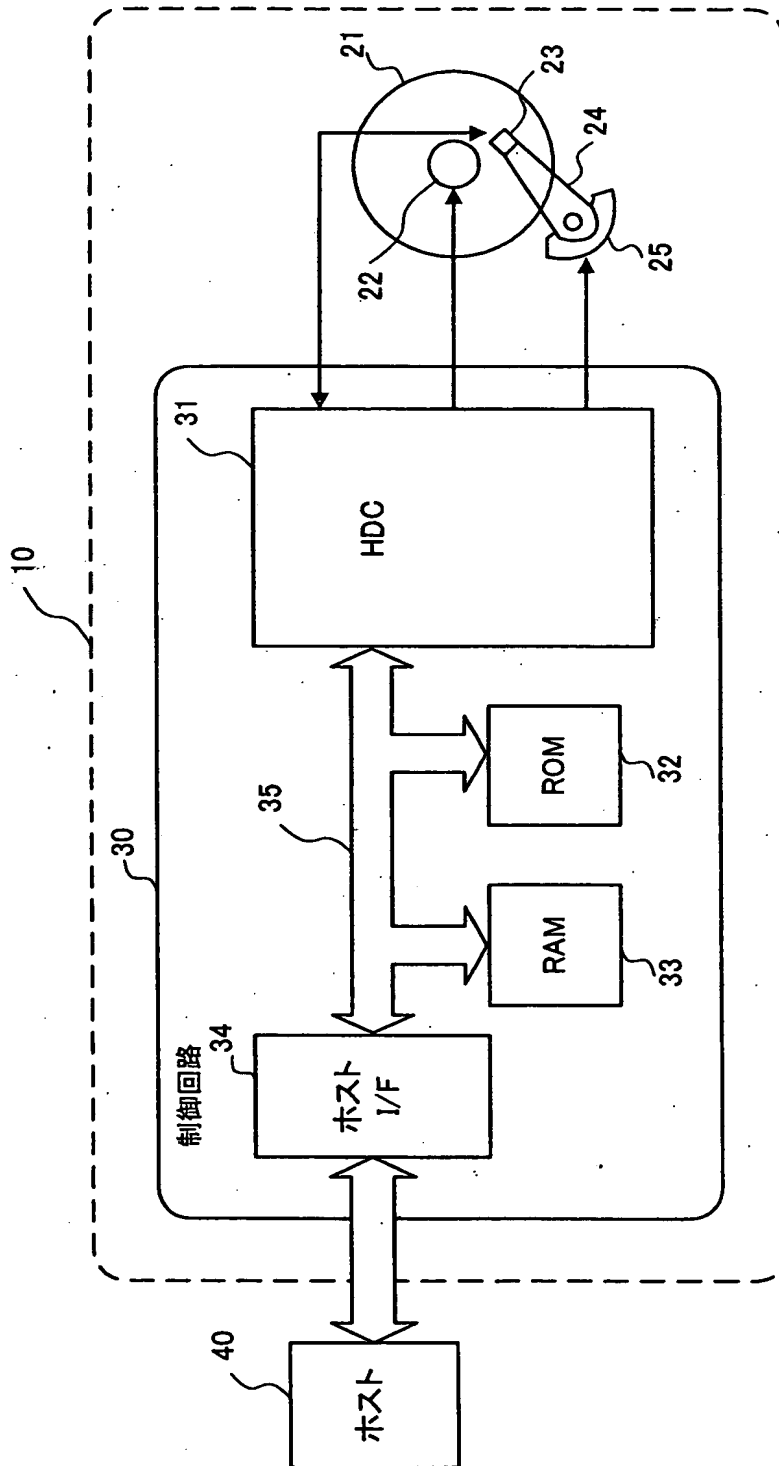


1 0 … HDD (Hard Disk Drive) 装置、 2 1 … 磁気ディスク、 2 2 … スピンドルモータ、 2 3 … 磁気ヘッド、 2 4 … ヘッドアーム、 2 5 … アクチュエータ、 3 0 … 制御回路、 3 1 … HDC (Hard Disk Controller)、 3 2 … ROM、 3 3 … RAM、 3 4 … ホスト I / F (インターフェース)、 3 5 … バス、 4 0 … ホスト、 4 1 … MPU (マイクロ・プロセッシング・ユニット)、 4 2 … MM (メモリ・マネージャ)、 4 3 … RAM、 4 4 … SG (セグメント・ハンドラ)。

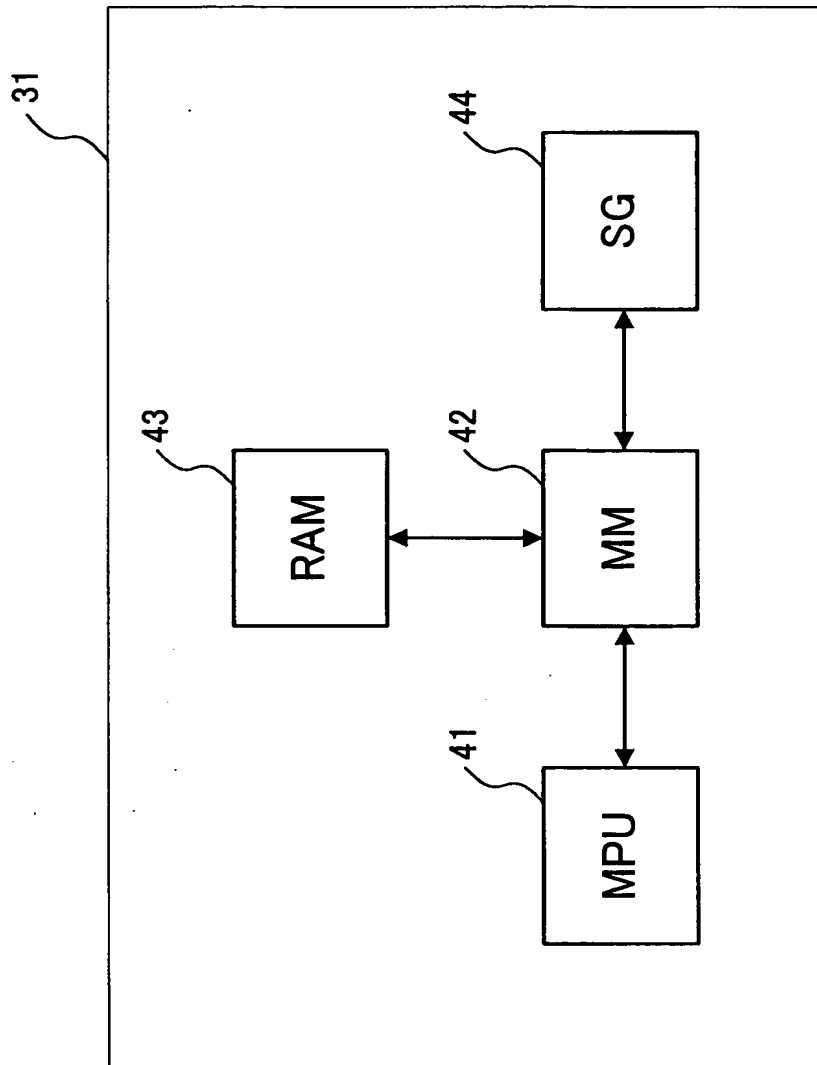
【書類名】

図面

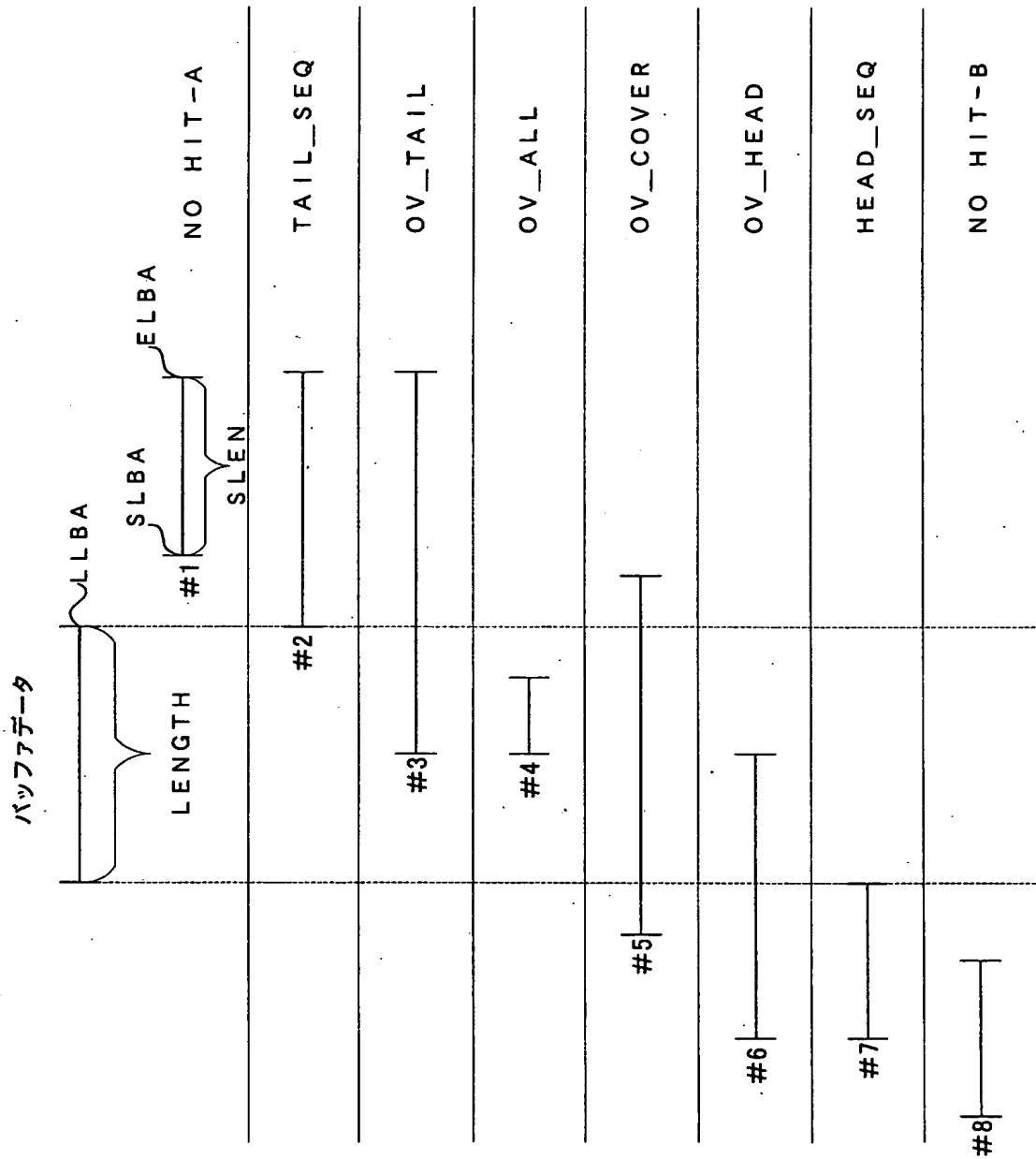
【図 1】



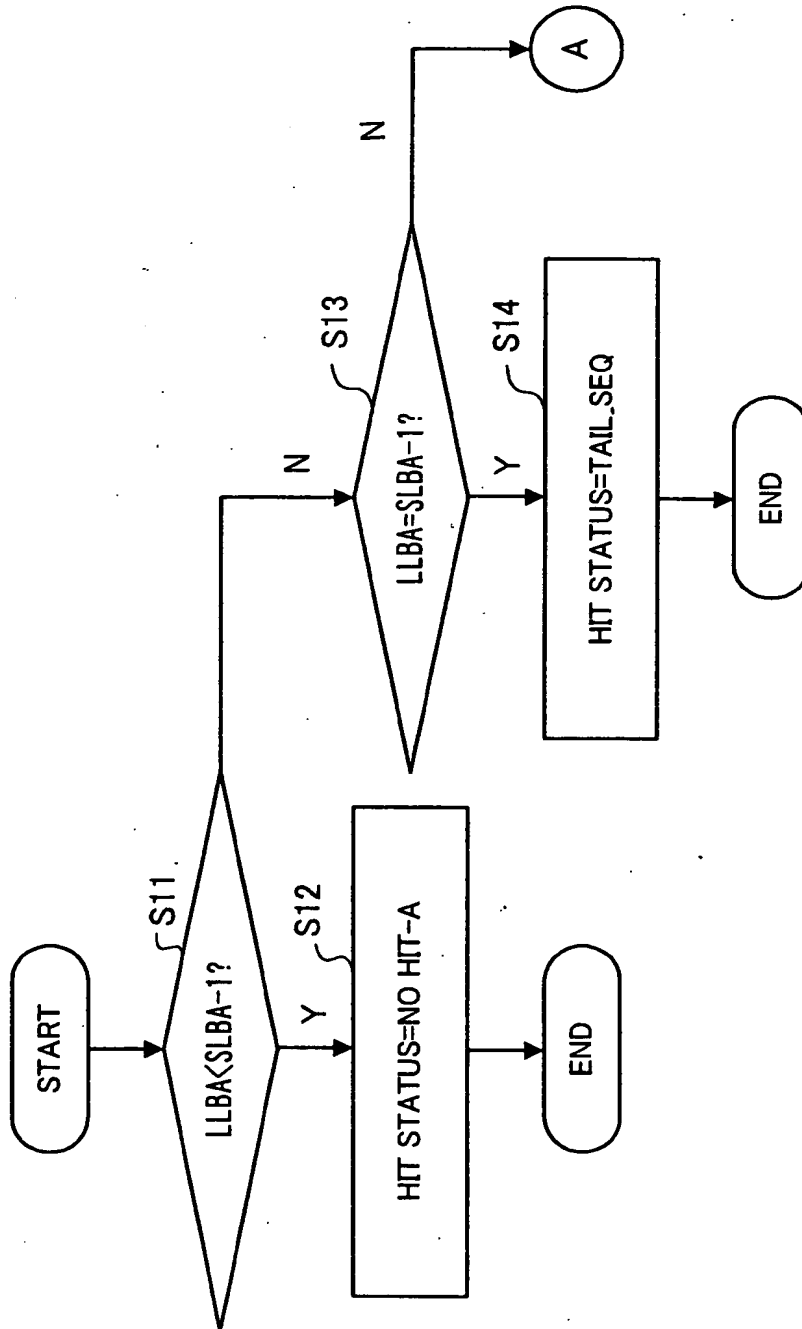
【図 2】



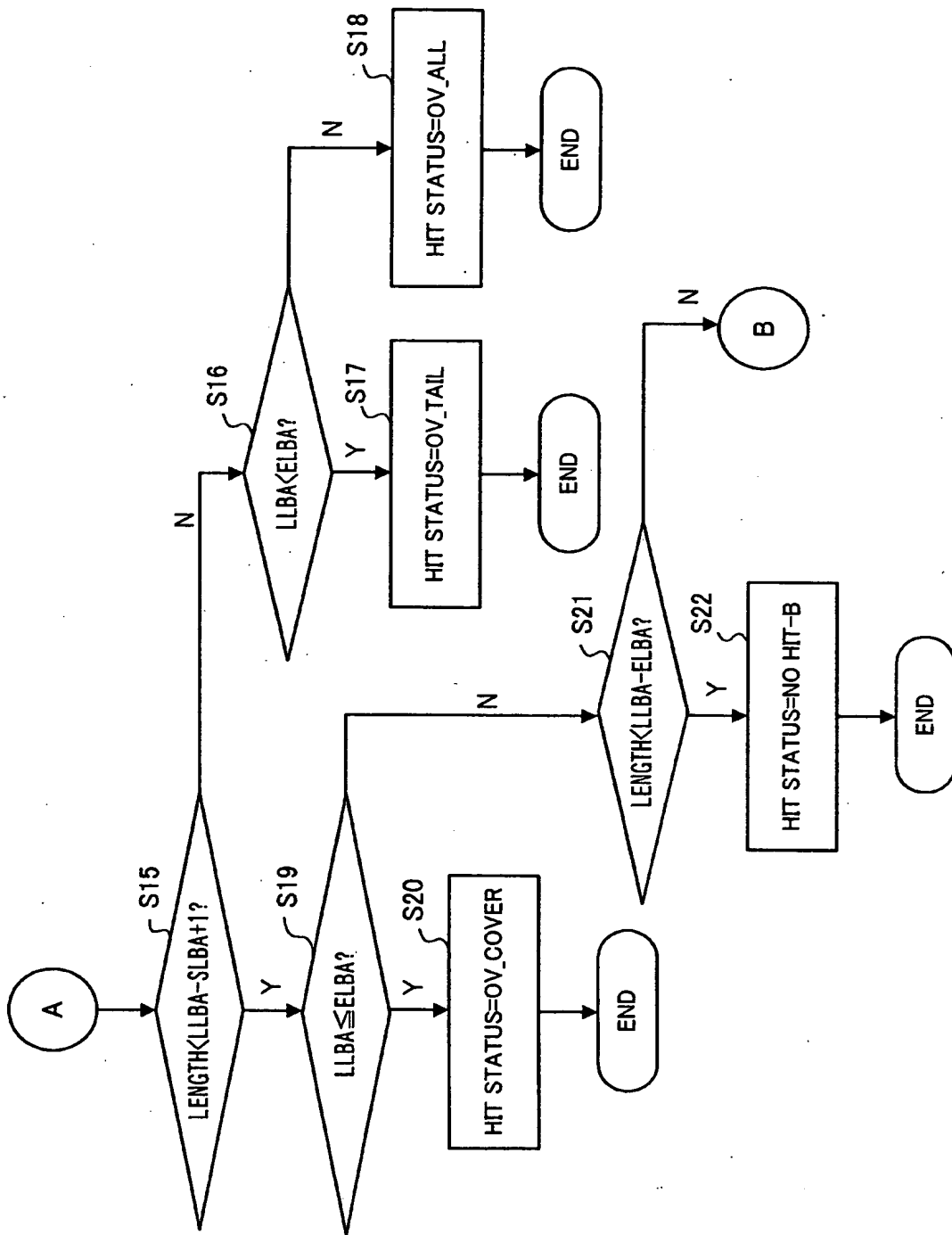
【図 3】



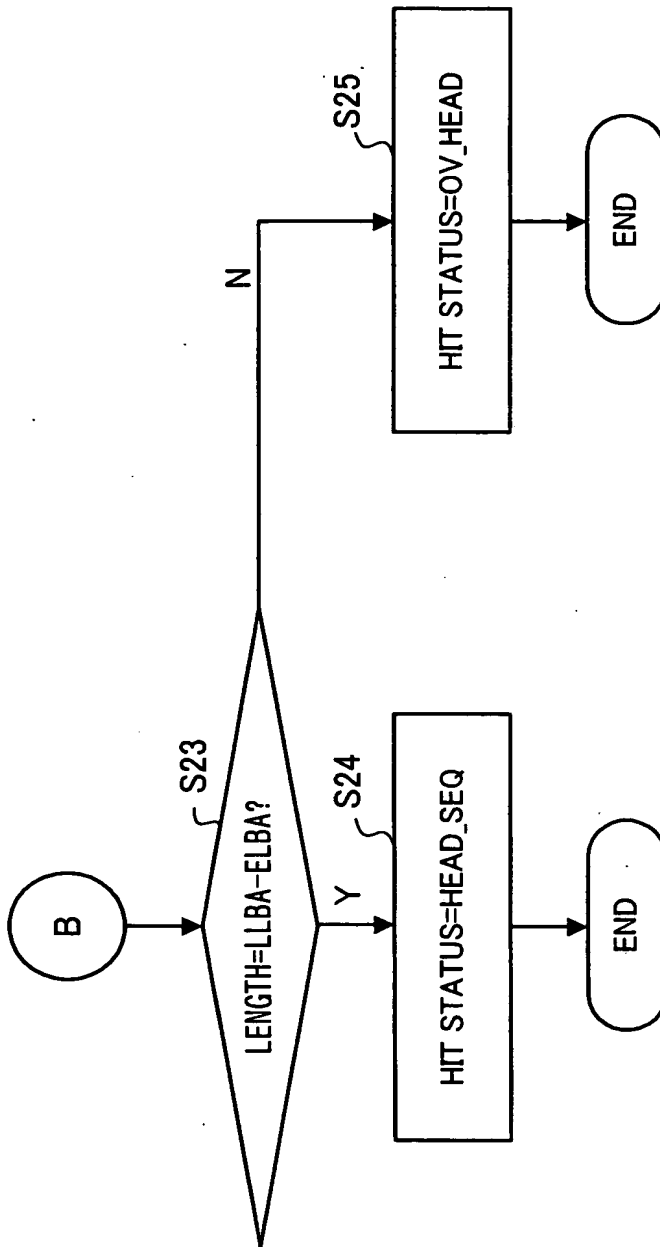
【図 4】



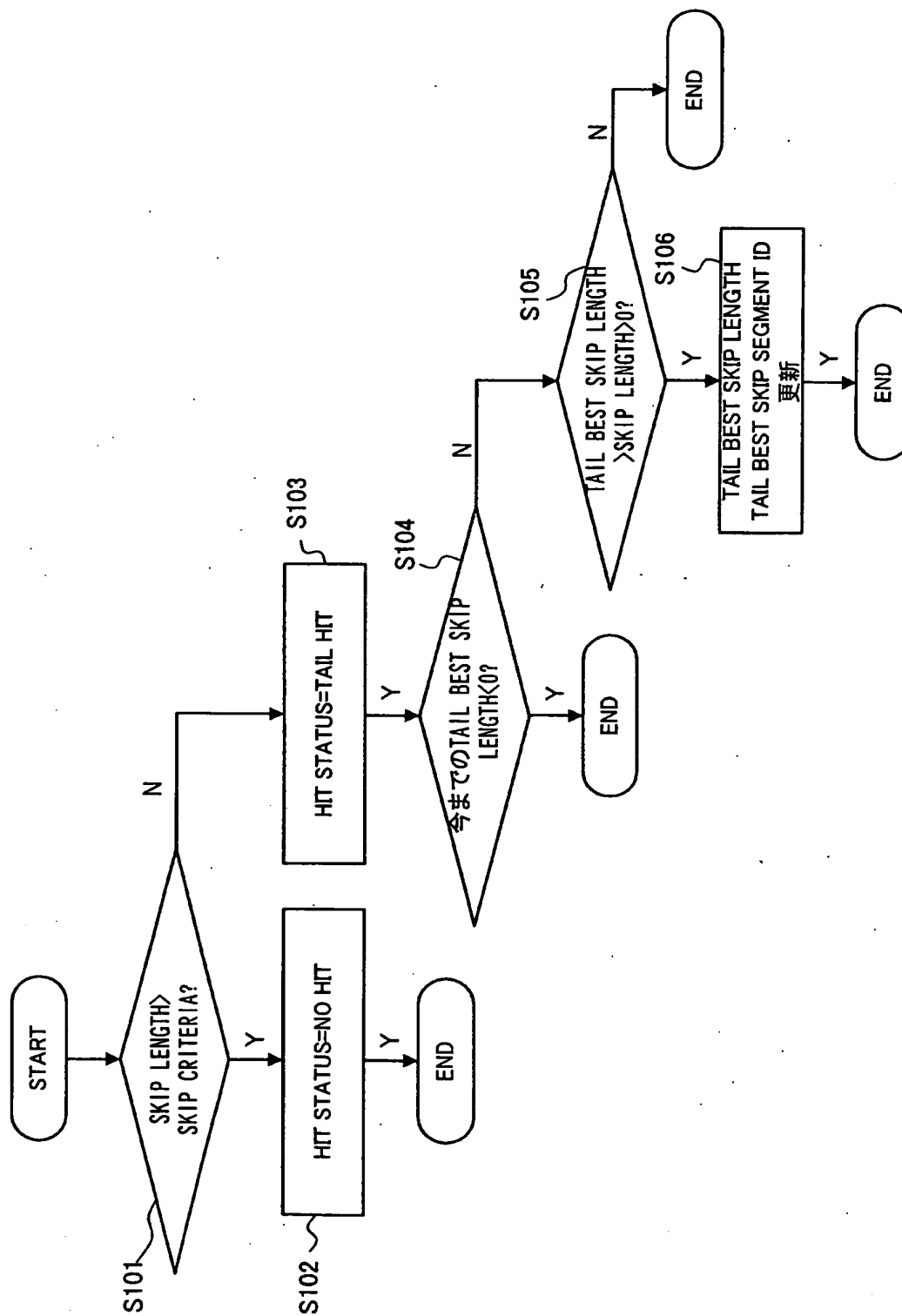
【図 5】



【図6】

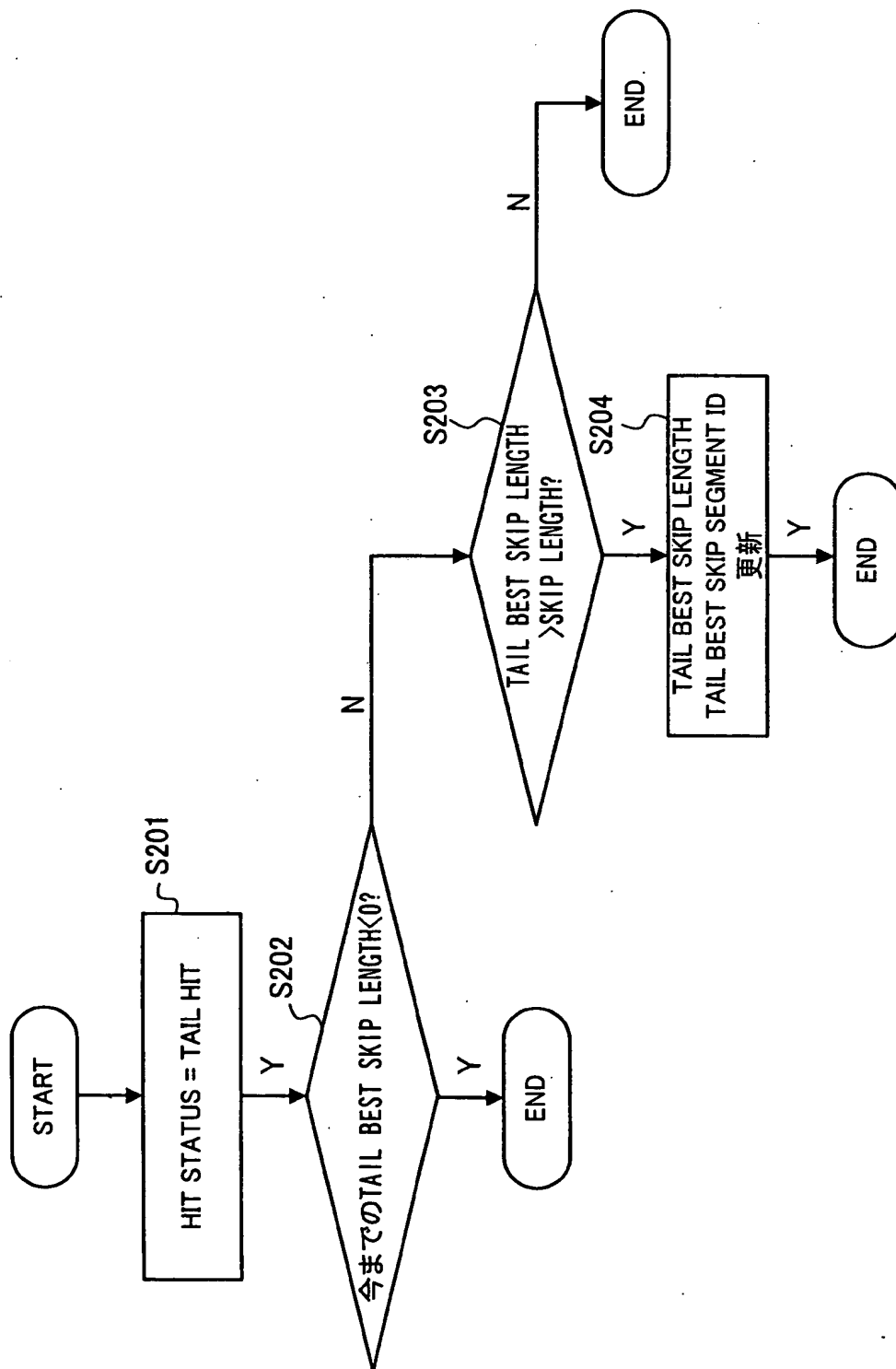


【図 7】

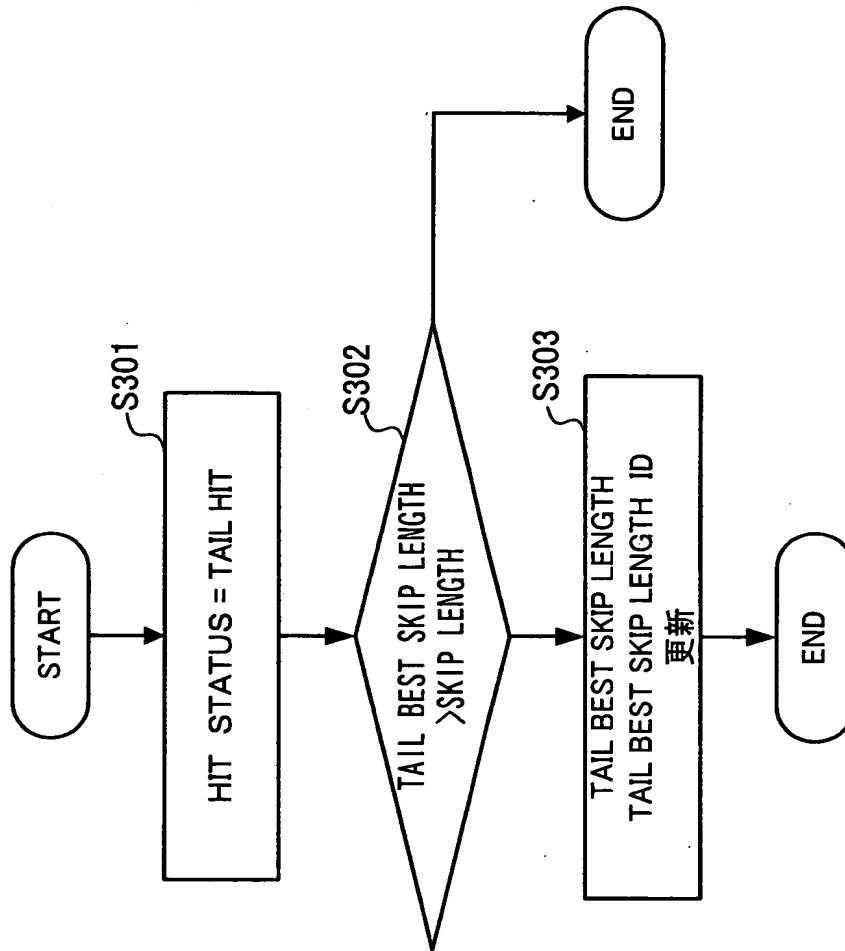




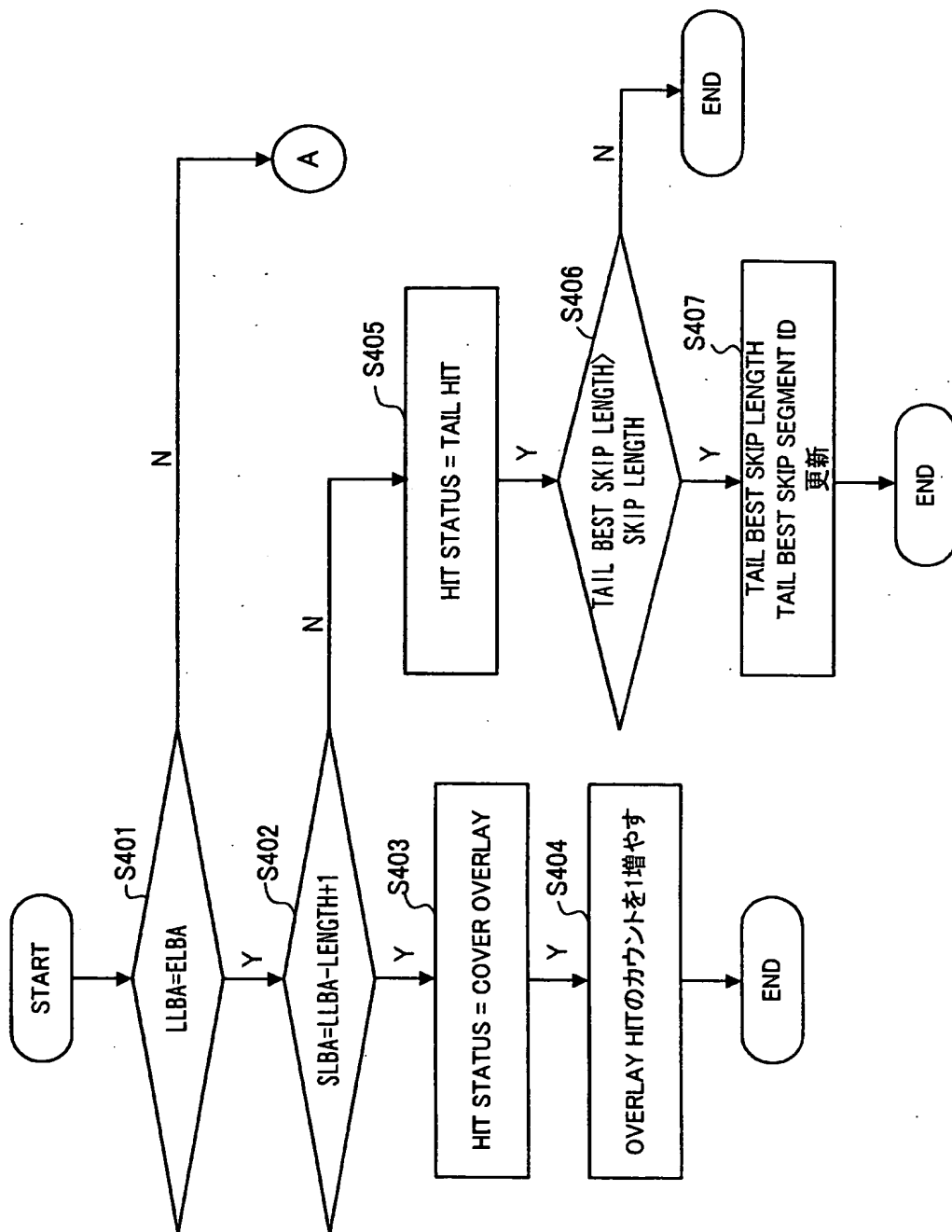
【図 8】



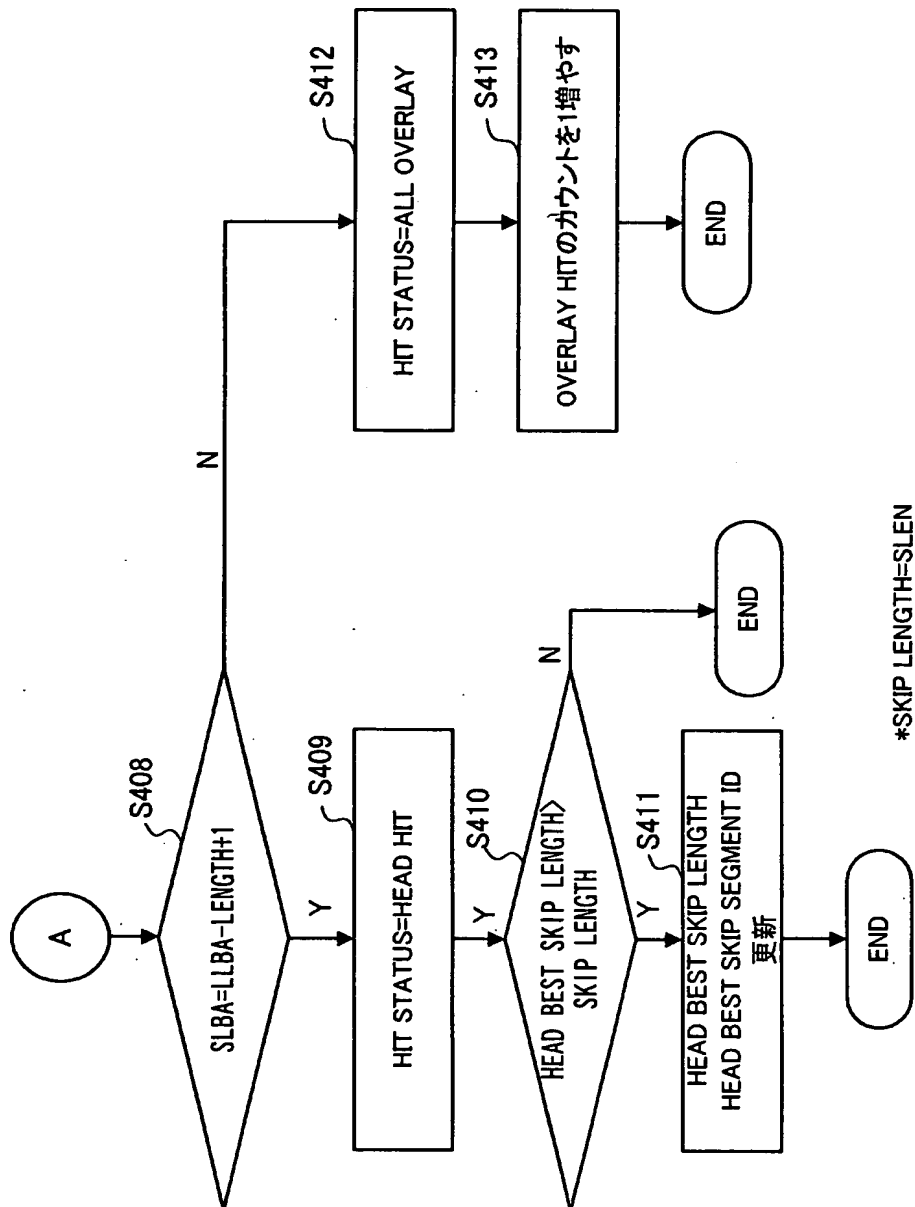
【図 9】



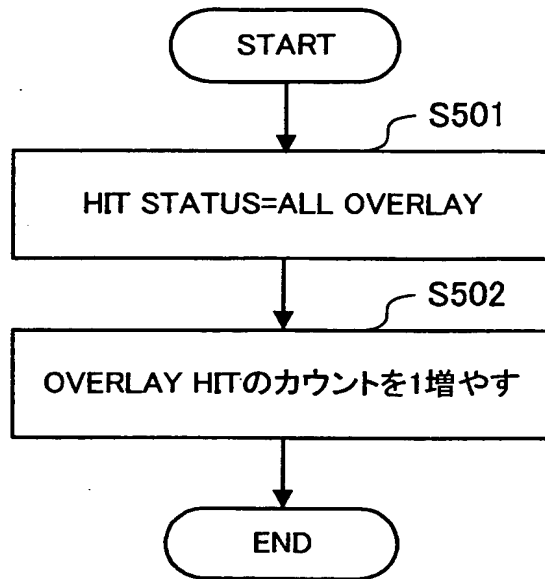
【図10】



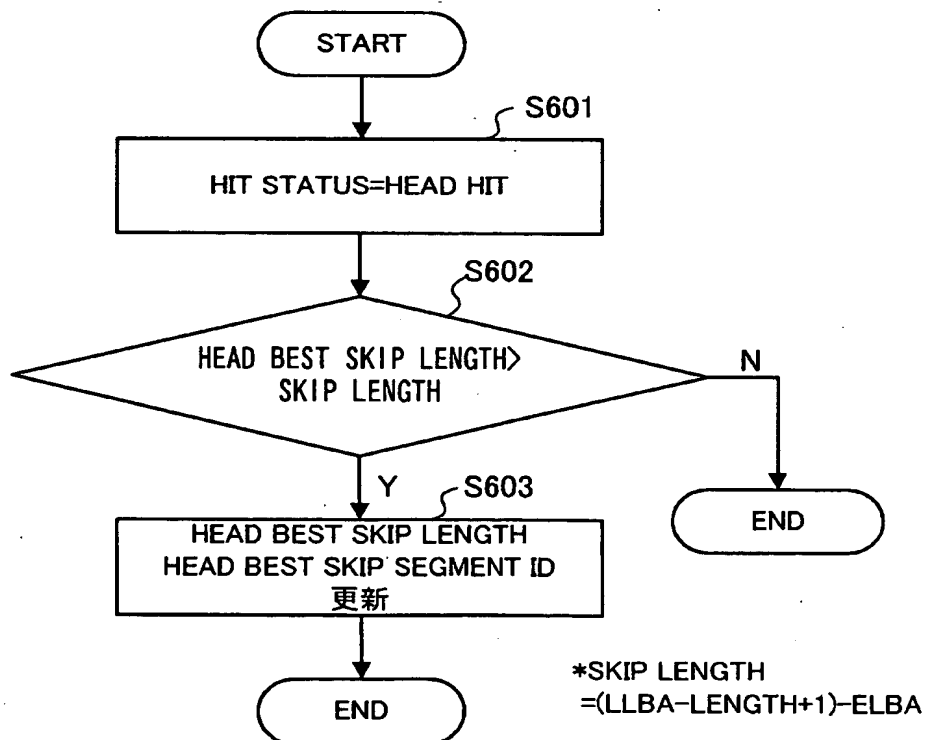
【図 11】



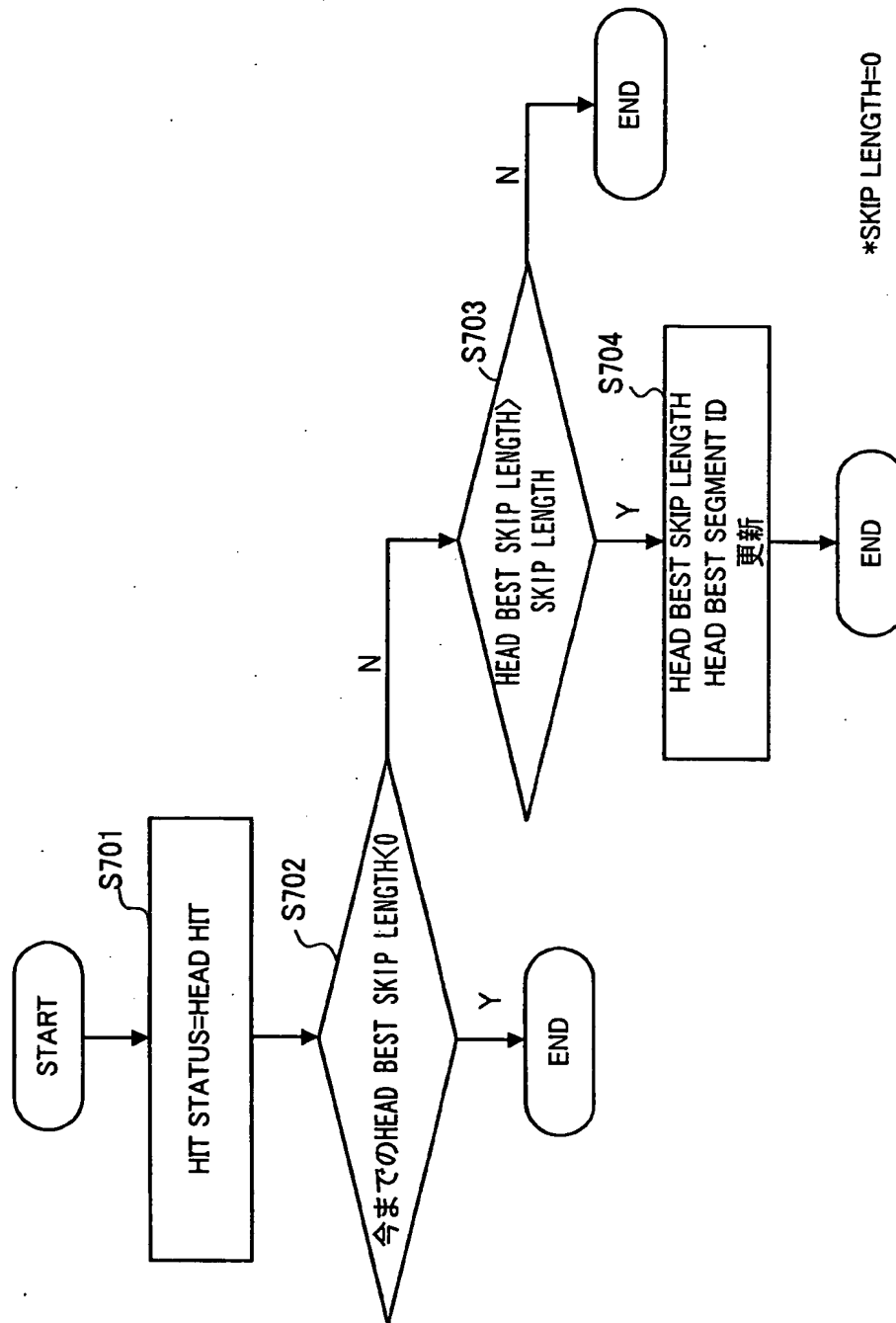
【図 1 2】



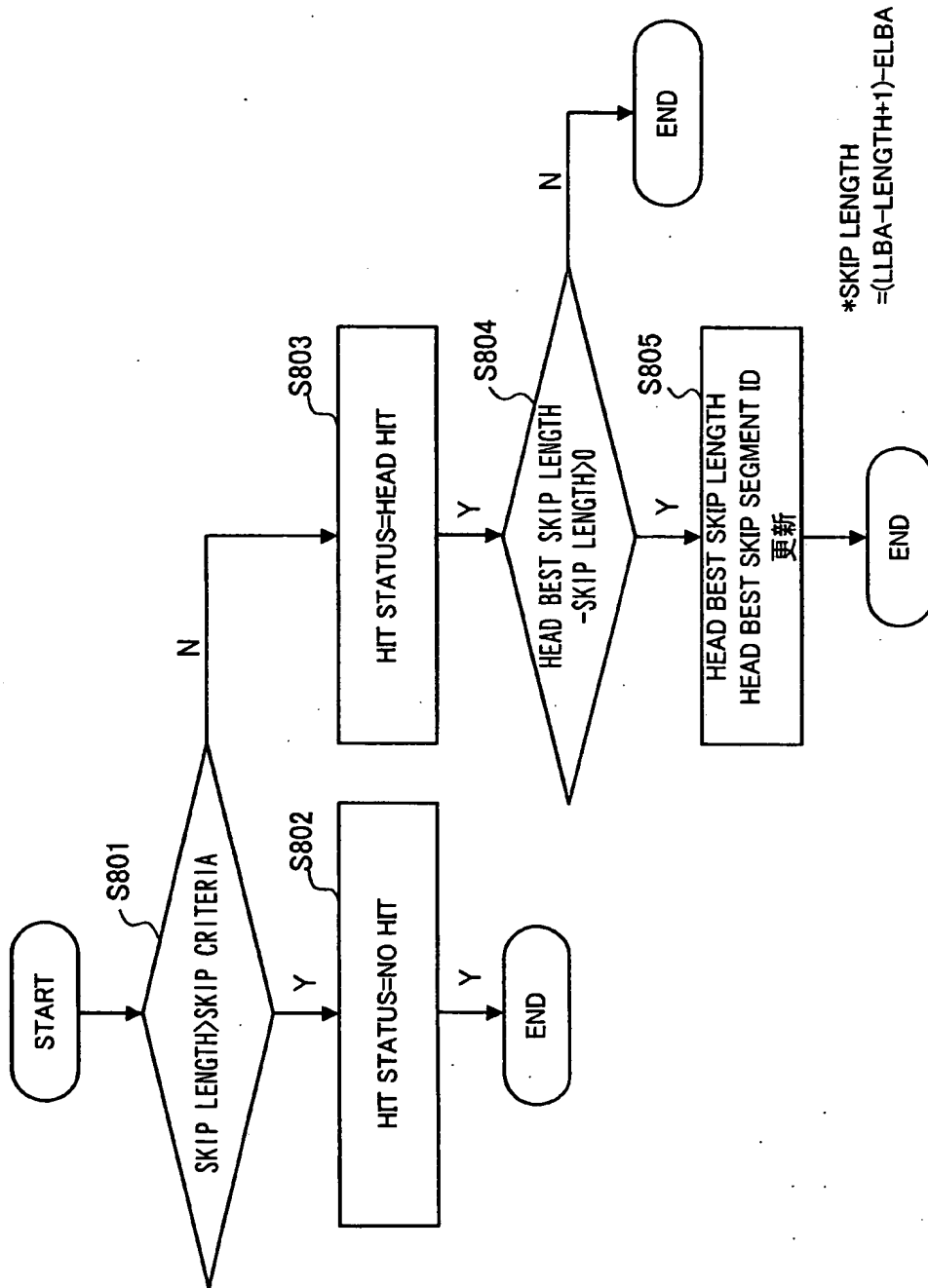
【図 1 3】



【図 14】



【図 15】



【図 16】

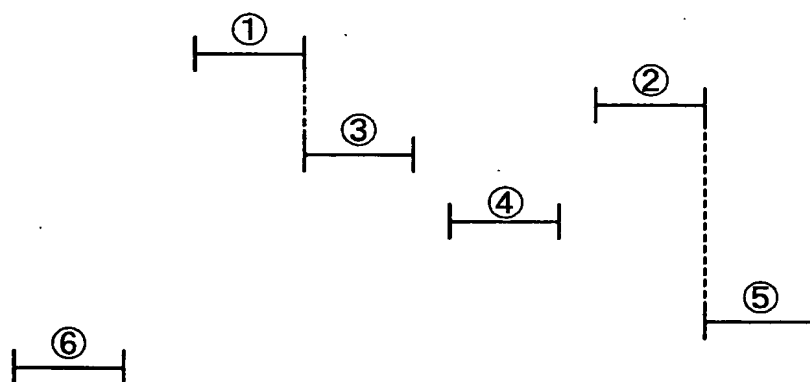
H/W USE	H/W USE	H/W USE	NO OF HIT
HEAD BEST SKIP LENGTH	HEAD BEST SKIP SEGMENT ID	TAIL BEST SKIP LENGTH	TAIL BEST SKIP SEGMENT ID

【図 17】

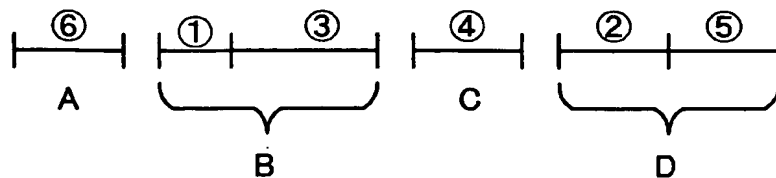
LAST LBA LOW	LAST LBA HIGH	LENGTH	HIT STATUS
--------------	---------------	--------	------------

【図 18】

(a)

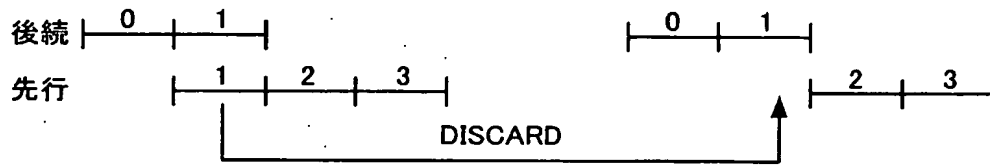


(b)

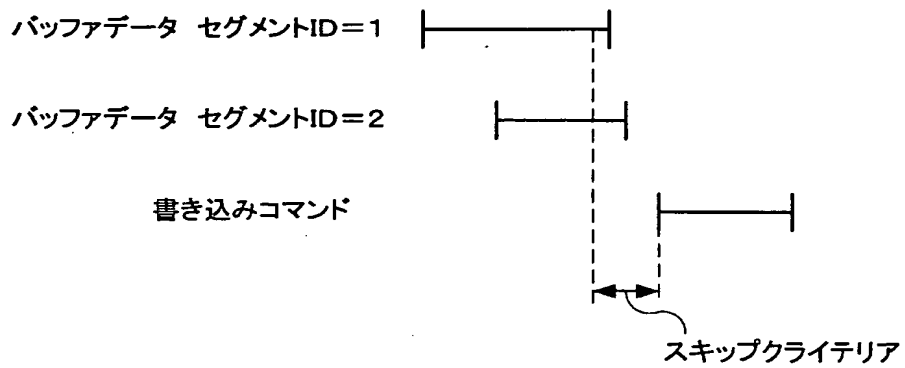




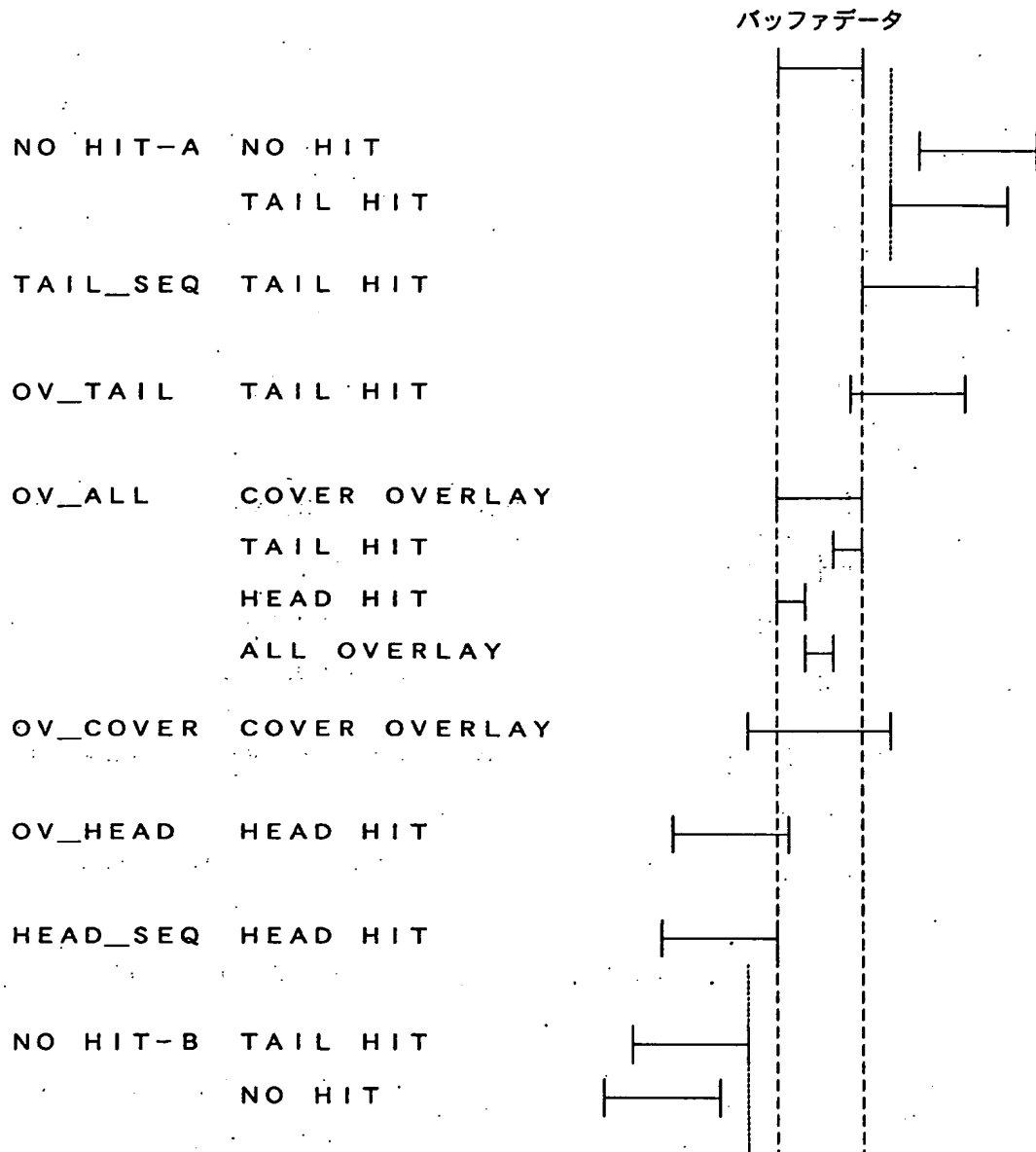
【図 1 9】



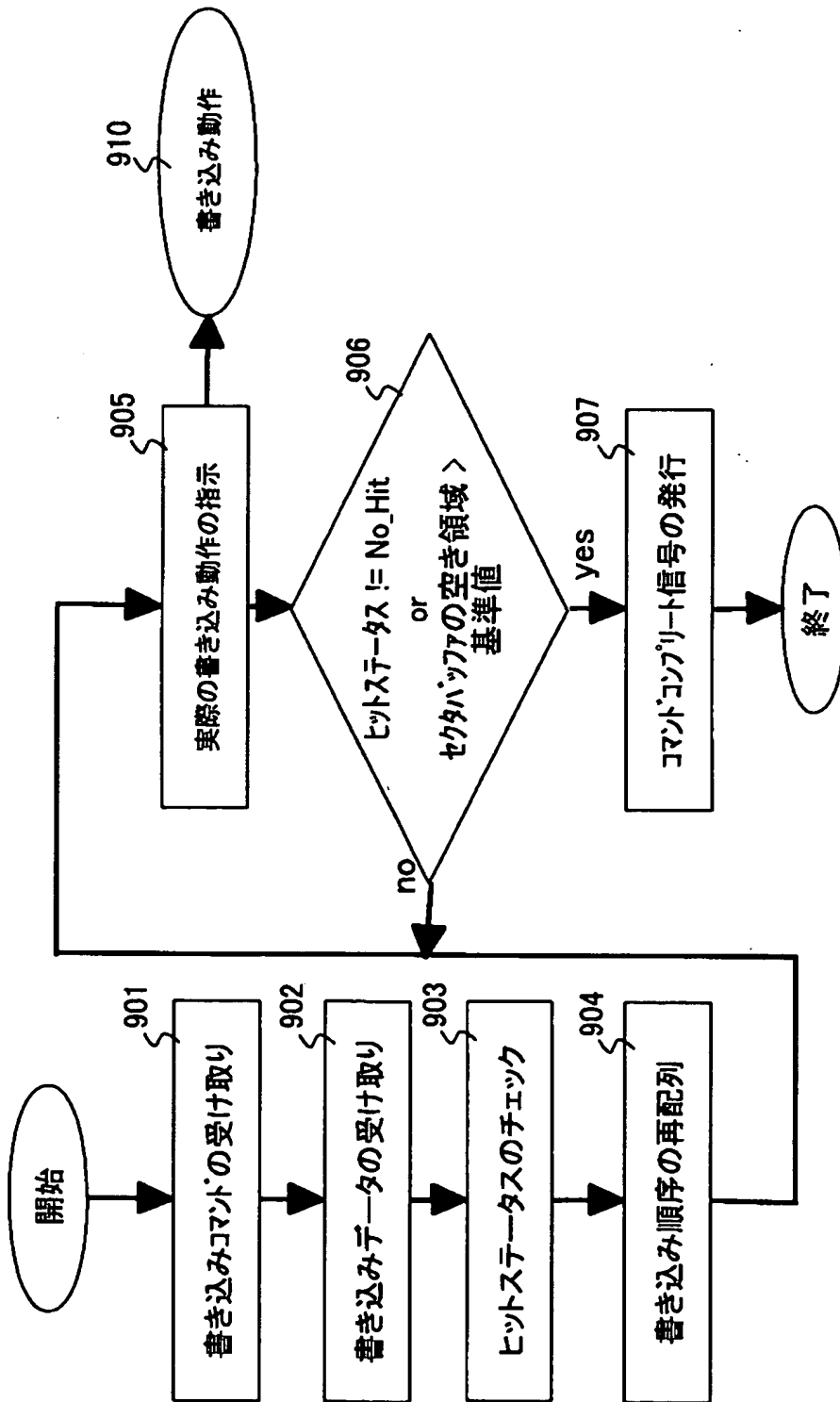
【図 2 0】



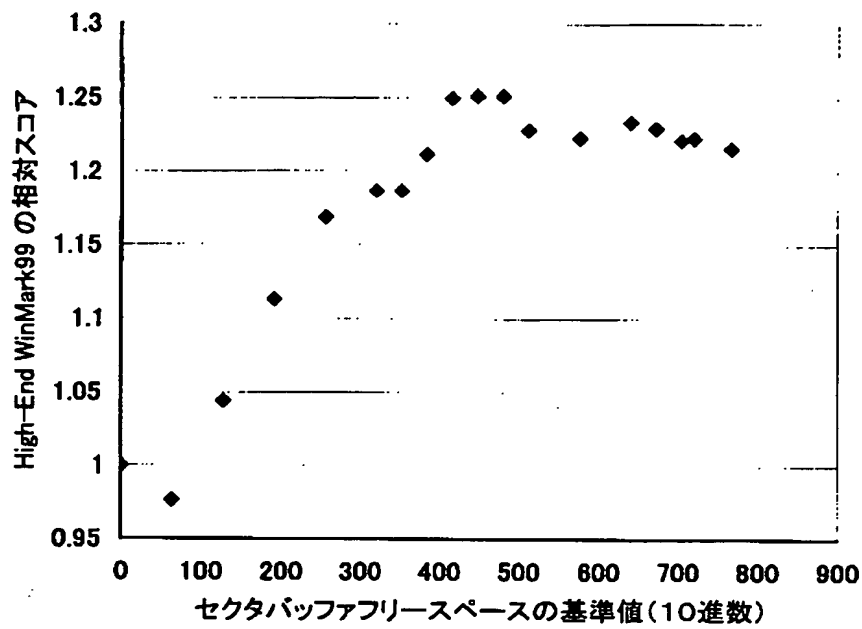
【図 21】



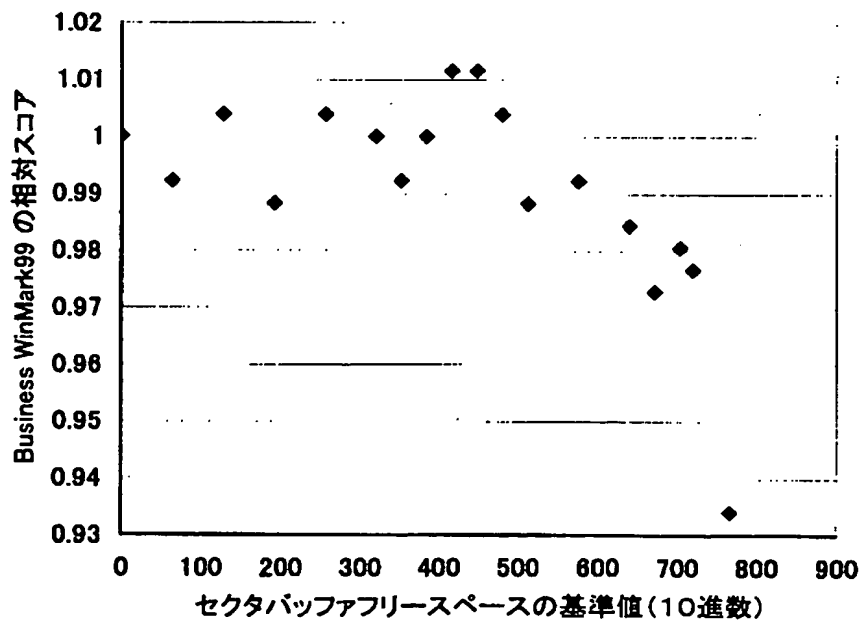
【図 22】



【図 2 3】



【図 2 4】



【書類名】                      要約書

【要約】

【課題】    セクタバッファが占有される状態を回避し、書き込み時の待ち時間を短縮してハードディスクドライブ装置のパフォーマンスを向上する。

【解決手段】    既保持の書き込みコマンドが実行待ち状態にあるとき、新たな書き込みコマンドを受け取り、既保持コマンドのLBAとの位置関係を表すヒットステータスをチェックし、記録媒体への書き込み順序を再配列し、再配列された書き込み順序に従って媒体に書き込みデータを記録する処理を行う。このとき、セクタバッファの空き領域が規定の基準値以下であり、かつヒットステータスがNO HITであるときには受け入れたコマンドに対するコマンドコンプリート信号を発行せず、媒体への記録処理を優先する。

【選択図】                      図 2 2

認定・付加情報

特許出願の番号	特願2000-121853
受付番号	50000510873
書類名	特許願
担当官	塩崎 博子 1606
作成日	平成12年 6月 5日

<認定情報・付加情報>

【特許出願人】

【識別番号】	390009531
【住所又は居所】	アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
【氏名又は名称】	インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

【識別番号】	100086243
【住所又は居所】	神奈川県大和市下鶴間1623番地14 日本アイ・ビー・エム株式会社 大和事業所内
【氏名又は名称】	坂口 博

【復代理人】

【識別番号】	申請人
【識別番号】	100112520
【住所又は居所】	神奈川県大和市中心林間3丁目4番4号 サクライビル4階 間山・林合同技術特許事務所
【氏名又は名称】	林 茂則

【選任した代理人】

【識別番号】	100091568
【住所又は居所】	神奈川県大和市下鶴間1623番地14 日本アイ・ビー・エム株式会社 大和事業所内
【氏名又は名称】	市位 嘉宏

【選任した代理人】

【識別番号】	100106699
【住所又は居所】	神奈川県大和市下鶴間1623番14 日本アイ・ビー・エム株式会社大和事業所内
【氏名又は名称】	渡部 弘道

【選任した復代理人】

【識別番号】	100110607
--------	-----------

認定・付加情報（続き）

【住所又は居所】	神奈川県大和市中央林間3丁目4番4号 サクラ イビル4階 間山・林合同技術特許事務所
【氏名又は名称】	間山 進也
【選任した復代理人】	
【識別番号】	100098121
【住所又は居所】	神奈川県大和市中央林間3丁目4番4号 サクラ イビル4階 間山・林合同技術特許事務所
【氏名又は名称】	間山 世津子

出 願 人 履 歴 情 報

識別番号 [390009531]

1. 変更年月日 1990年10月24日  
[変更理由] 新規登録  
住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)  
氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション
  
2. 変更年月日 2000年 5月16日  
[変更理由] 名称変更  
住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)  
氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション